UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS

COLEGIO DE FILOSOFIÁ

"APLICACIONES DE LOS MÉTODOS DE PROBABILIDAD BAYESIANO Y FRECUENCIALISTA EN EL MÉTODO CIENTÍFICO PARA CONFIRMAR HIPÓTESIS CIENTÍFICAS"

TESIS QUE PARA OBTENER EL TÍTULO DE
LICENCIADA EN FILOSOFÍA
PRESENTA:

NAYELI RODRÍGUEZ DE JESÚS

DIRECTORES DE TESIS:

MTRA. GABRIELA HERNÁNDEZ DECIDERIO.

DR. FEDERICO MARULANDA REY.

MÉXICO D.F. ABRIL DE 2011

Agradecimientos

En esta parte de la tesis siempre uno se convierte en injusto porque se olvida, por no tener registro escrito, de las muchas personas que contribuyeron a que este trabajo tenga refinamiento en la escritura, la investigación, la gesta del proyecto en sí, el desarrollo de la idea. Enumerar a las personas que me ayudaron sería interminable, ya que por suerte cuento con gente que me ayuda, me ayudó y espero contar con su apoyo como ellos de mí siempre lo tendrán.

Una de las personas a la que estoy muy agradecida es la que estuvo conmigo al inicio de esta tesis, lamentablemente no pudo estar conmigo al final, sin embargo fue esencial su apoyo y dedicación, ya que gracias a él me atreví a adentrarme a una disciplina muy poco desarrollada en el ámbito académico. Al igual agradezco a la persona que lo ha apoyado siempre. Nunca les podré agradecer lo suficiente el enorme esfuerzo de soportarme durante tanto tiempo y con mis pocos conocimientos de la disciplina.

A las personas que me soportaron y tuvieron paciencia conmigo, en todo este proceso. En particular, Agradezco a la Dra. Lourdes Valdivia, a la Mtra. Gabriela Hernández, al Dr. Federico Marulanda, al Dr. Ricardo Vázquez y al Lic. Cristian Gutiérrez. Gracias por su apoyo cuando más lo necesitaba, pues esta tesis no hubiera salido a flote sin sus críticas, comentarios. Por último, y no menos importante, a aquellas personas pacientes y tolerantes que estuvieron conmigo en este proceso final de la tesis, en particular agradezco al Dr. Raymundo Morado por su apoyo incondicional. Dedico esta investigación a las personas que me han apoyado toda la vida.

ÍNDICE

Introducción

En el proceso de investigación de las ciencias experimentales, para poder discriminar entre hipótesis científicas, es decir, corroborar o rechazar hipótesis dada la evidencia, no es suficiente apoyarse en un método científico deductivo, ya que en este tipo de investigaciones intervienen aspectos observacionales que nos exigen la introducción de un método probabilístico al científico deductivo, para seleccionar de una manera más rigurosa y formal las hipótesis científicas. Lo anterior da origen a preguntas de relevancia para la selección de hipótesis que se apoyan en la evidencia, preguntas como: qué dato debe ser considerado como evidencia, qué evidencia es relevante para apoyar la hipótesis *x*.

En el presente trabajo nos interesa destacar dos métodos o enfoques de la probabilidad que dan respuestas a dichas preguntas, los dos métodos son: el método frecuencialista y el método bayesiano. El primero tiene una postura objetiva de la probabilidad y el segundo tiene una postura subjetiva de la probabilidad[1]. Los resultados que se obtienen de cada uno pueden ser muy distintos, pues tienen diferentes posturas con respecto a la forma de calcular la probabilidad de una hipótesis dada la evidencia. Por lo anterior, ahora surge nuestra pregunta, a saber, ¿qué método de la probabilidad deben aplicar los científicos, el método bayesiano o el frecuencialista?

Algunos filósofos como Wesley Salmon dicen que la respuesta a la pregunta anterior depende de la experiencia del científico, pues si obtuvo buenos resultados al aplicar tal método de la probabilidad en la corroboración o rechazo de alguna hipótesis, entonces lo

---

[1] La postura subjetiva apoya la teoría subjetiva de la probabilidad que es caracterizada por tratar con datos que tiene características cualitativas que afectan en el cálculo de la probabilidad de una hipótesis dada la evidencia; en cambio la postura objetiva apoya la teoría objetiva de la probabilidad que es caracterizada por tratar con datos que sólo tienen características cuantitativas que afectan en el cálculo de la probabilidad de una hipótesis dada la evidencia.

seguirá usando. Otros, como Bertrand Russell, buscan razones y factores que pueden ser claves para elegir alguno de los métodos de la probabilidad. Coincidimos con Russell en pensar que merece la pena identificar factores claves más allá de la experiencia previa para decidir cuál de los dos métodos de probabilidad, bayesiano o frecuencialista, tiene más ventajas. Ahora bien, tales factores pueden ser varios, pero en esta investigación nos basaremos sólo en dos de ellos que tienen que ver con las siguientes preguntas: ¿Qué método científico general se va a aplicar? ¿Qué problemas soluciona el método de la probabilidad empleado en el seno de dicho método? Dichos cuestionamientos serán fundamentales para dar razones a favor de las ventajas o desventajas del uso de alguno de los dos métodos de probabilidad. En este trabajo se expondrán las ventajas y desventajas del uso de los dos métodos de probabilidad, mencionados en los métodos científicos.

La primera pregunta, a saber, a qué método científico general se va a aplicar el método de la probabilidad, se abordará exponiendo el método científico que tenga la forma de evaluar una hipótesis de manera causal. Se expondrá este tipo de método porque se considera que es más claro para exponer los usos de los métodos de probabilidad[2]. La segunda pregunta, a saber, qué problemas soluciona o trata el método de la probabilidad que se va a usar para corroborar la hipótesis, también se expondrá sólo tratando algunos de los problemas fundamentales que tiene el método científico que pretende elucidar relaciones de causa-efecto. El método científico paradigmático para elucidar la estructura

[2] Voy a destacar el método hipotético deductivo de los tantos métodos que hay porque no voy a ver todos los métodos científicos ya que, este trabajo tiene la intención de ser específico sólo cuando trate sobre la confirmación de hipótesis en los métodos científicos, entre los que se destacaron están el método de la confirmación, el de la falsificación y el método hipotético deductivo científico. Esto se hará para dar un contraste entre estos métodos. También se destacará el método hipotético deductivo científico, pues este método usa tanto el método frecuencialista como el bayesiano de la probabilidad, para apoyar la confirmación de hipótesis científicas.

causa-efecto es el método hipotético deductivo en el capítulo 3 expondremos dos de sus problemas fundamentales: el problema de la contrastación de hipótesis y el problema de la observación de la evidencia.

La pregunta principal que enfrentará esta tesis es: ¿qué método de la probabilidad, bayesiano o frecuencialista, tiene más ventajas cuando se usa en el método científico hipotético deductivo para corroborar hipótesis? La hipótesis principal de nuestra tesis es que en este contexto el método bayesiano tiene más ventajas que el método frecuencialista de la probabilidad para corroborar hipótesis. Para sustentar nuestra hipótesis principal tenemos que exponer los problemas fundamentales que tienen los métodos científicos con la estructura causa-efecto y que pueden ser solucionados o, por lo menos, tratados por alguno de los dos métodos de probabilidad ya señalados.

Ahora bien, uno de los problemas que enfrenta el método científico es la contrastación de hipótesis y sus consecuencias, este problema será desarrollado en la capítulo 3. Si bien contamos con una gran variedad de métodos científicos, cuando se aplica un método científico a un evento x que corrobora una hipótesis y, puede ocurrir que otro método científico arroje un resultado diferente sobre la hipótesis y referente al mismo evento x[3]. Ante tal situación de discrepancia, los científicos pueden buscar un criterio por medio del cual puedan apoyar su corroboración o rechazo de la hipótesis y sobre el evento x; en este caso, por lo general se preguntan si la hipótesis es adecuadamente apoyada por la

[3] Por ejemplo, si nosotros usamos el método científico de falsificación para evaluar la hipótesis de que el agua de tlacote cura el dolor de cabeza; con sólo tener un caso desconfirmador sobre la hipótesis, es suficiente para desvirtuarla. En cambio, si usamos el método científico de la confirmación sólo necesitamos un caso confirmador de la hipótesis para confirmarla. Si usamos el método hipotético deductivo científico, la hipótesis debe de tener varios casos corroboradores y debe tener la probabilidad más alta de las hipótesis alternativas.

evidencia; esta pregunta genera otra, a saber, ¿cuándo podemos decir que una hipótesis *x* es adecuadamente apoyada por la evidencia?

Si bien una hipótesis científica puede ser corroborada o invalidada por la evidencia que obtengamos de la observación del evento *x,* en las varias pruebas que se realizan a la hipótesis, surgen varias dificultades en este proceso. Una es cuando hablamos de observación, pues en ella están implicadas la subjetividad y distintas perspectivas, por ende, también distintas opiniones de la observación del evento. De esta forma, podemos tener varias hipótesis científicas distintas que pueden ser corroboradas por la observación del mismo evento *x*. Ante tales problemas, los científicos recurrieron a un método que les pudiera proporcionar rigor y formalismo, que se pudiera tratar con argumentos inductivos, además de que fuera independiente de los problemas que implica la observación. Tal método es, en principio, aportado por la probabilidad: es decir, un razonamiento probabilístico debería permitir la selección de la hipótesis más probable entre varias alternativas, lo que resulta más riguroso elegir una hipótesis de entre varias únicamente mediante la observación.

Aunque la lógica deductiva es un apoyo primordial y formal en los métodos científicos para la selección de hipótesis; la probabilidad expone de una manera clara y formal las hipótesis alternativas que se tienen, asignando una probabilidad a cada hipótesis por medio de una fórmula de probabilidad, ya sea del método bayesiano, frecuencialista o algún otro.

Podemos decir que entre más experimentos se realicen a las hipótesis alternativas, más notoria será la diferencia de qué hipótesis resalta o es más plausible, sin embargo, si el científico quiere que sus experimentos sean considerados como serios y rigurosos, no basta con que los experimentos que realice den como resultado hipótesis plausibles. Lo que necesita el científico es que sean muy probables, ya que la probabilidad le dará un respaldo de rigor a los casos que corroboren la hipótesis, pues, si nos preguntamos ¿cuáles son las probabilidades de que la predicción inferida de las observaciones del evento *x* fuera cierta, si la hipótesis que estamos verificando fuera falsa y alguna otra fuera cierta? Es decir ¿existen otras hipótesis que resultarían fuertemente confirmadas por el mismo resultado?

Supóngase que se ha visto que el agua de tlacote cura el dolor de cabeza, la hipótesis es que tomar agua de tlacote cura el dolor de cabeza. Las condiciones iniciales son que se tiene dolor de cabeza y se toma agua de tlacote. La predicción de observación es que desaparecerá el dolor de cabeza. El dolor de cabeza desaparece, efectivamente, y esto constituye un caso confirmador o corroborador de la hipótesis; se tienen varios casos confirmadores de la hipótesis[4], es decir, una hipótesis puede tener casos confirmadores cuando se realizan varios experimentos sobre un evento donde nos interesa saber si ocurre o no, pues si ocurre en un experimento se considerará como un caso que confirma la hipótesis. Debemos preguntarnos ¿ha sido la hipótesis confirmada por estos casos confirmadores? Digamos que hay una hipótesis alternativa que sugiere que el dolor de cabeza es un síntoma psicosomático que puede curarse por sugestión, es decir, cualquier tratamiento que el paciente crea sinceramente que es efectivo lo será. Sabido esta hipótesis, la pregunta sería ¿qué hipótesis debe ser confirmada? Si la última hipótesis alternativa es

[4] Cuando decimos que una hipótesis puede ser fuertemente confirmada nos referimos a que podrían tener el mismo caso de casos confirmadores que la hipótesis. Por ejemplo, en una comunidad se cree que cierta hierba quita el dolor de cabeza y es fuertemente confirmada su hipótesis, porque tiene varios casos que la confirman, tanto como los casos que confirman que el agua de tlacote quita el dolor de cabeza.

correcta, el frotar una hierba en la frente sería igualmente eficaz a condición de que el paciente crea que se va a curar.

La probabilidad que hay de las anteriores hipótesis es de relevancia, pues pueden desvirtuar tanto a la primera como a la segunda hipótesis y lo que hace la probabilidad es aclarar las hipótesis científicas asignando una probabilidad, pues si ésta no se usara, no habría forma de diferenciar las distintas características que son relevantes en el conteo de los casos confirmadores de la hipótesis científica. Distinguir las distintas características del evento que se presentan en los casos confirmadores, afectan considerablemente a los resultados. La probabilidad se encarga de considerar ese tipo de características que se atribuyen a los eventos; además de considerar los distintos tipos de experimentos que se pueden hacer, por ejemplo, hace la distinción de experimentos compuestos y simples. Este tipo de consideraciones son relevantes para obtener una probabilidad de la hipótesis que no sea errónea o se apele a una falacia de probabilidad; también, trata de ver que los métodos de probabilidad que se apliquen cumplan con los axiomas de la teoría axiomática de la probabilidad, pues no cualquier teoría de la probabilidad puede ser legítima, por lo menos debe de cumplir con dichos axiomas, para ser considerada como legítima.

Según lo expuesto anteriormente, hemos explicado las ventajas de los métodos de probabilidad en los métodos científicos, ahora podemos decir que de los varios métodos de probabilidad que hay, en este texto, sólo nos vamos a enfocarnos en los métodos bayesiano y frecuencialista de la probabilidad, pues por lo general son los más usados y se desglosan de dos posturas distintas de la probabilidad, a saber, la probabilidad cualitativa que corresponde al método bayesiano y la probabilidad cuantitativa que corresponde al método frecuencialista. Estas dos posturas de la probabilidad fueron evolucionando hasta convertirse en las posturas dominantes en el ámbito de los métodos de probabilidad que son usados en los métodos científicos para confirmar hipótesis.

A continuación, las ventajas del método bayesiano:

- Tiene una estructura causal y su estructura formal le permite tener esta ventaja.
- Soluciona el problema de la probabilidad inversa, que afecta de manera positiva en la asignación de probabilidad a las hipótesis.
- Puede cambiar su probabilidad inicial cuando obtiene más información, es decir, ésta no es fija.
- Tiene la facilidad de ser aceptada, tanto por la interpretación objetiva de la probabilidad como por la interpretación subjetiva, es decir, no se restringe su uso.
- Trata con evidencia que puede tener características tanto cuantitativas como cualitativas.
- Puede manejar eventos dependientes, independientes, compuestos y simples.
- Su estructura formal le permite tratar con distintos tipos de eventos, ya que implica varias reglas de probabilidad, como pueden ser: la regla del producto y la de la suma de la probabilidad que son usadas en el método bayesiano para calcular la probabilidad total de eventos compuestos.
- Trata el problema de la probabilidad *a priori* o inicial, pues propone varias soluciones para determinarla, una de ellas es el cálculo de la coherencia.

Sin embargo, el método de Bayes, también tiene desventajas, a saber:

- Aún no se tiene una manera formal y rigurosa para determinar la probabilidad *a priori* que sea aceptada como legítima.
- Pierde formalidad cuando maneja eventos con características cualitativas, pues el cálculo de la probabilidad de su formalidad y rigor, no sólo se debe a los métodos o reglas formales que se usan, sino también a qué se están aplicando y por lo general, se aplican sólo a eventos cuantitativos. Esta restricción permite que tengan rigor, pues cuando no manejan eventos cualitativos se evita la ambigüedad que puede haber[5].

Por otra parte, el método frecuencialista de la probabilidad tiene las siguientes ventajas:

- Es formal y riguroso, pues trata con eventos cuantitativos que se pueden repetir.
- Tiene la estructura formal de causalidad, que usan por lo general los métodos científicos.

Sin embargo, tiene las desventajas de que:

- Su campo de estudio se limita a sólo eventos cuantitativos y a eventos que tienen frecuencias.
- Esta probabilidad sólo puede ser aceptada por la interpretación objetiva, de esta forma restringe su campo de utilidad, pues sólo trata con evidencia con características cuantitativas.
- No resuelve y mucho menos trata el problema de la probabilidad inversa, por lo que se ve afectado en su uso para los métodos científicos, entre ellos el método hipotético deductivo científico.

[5] Por ejemplo, dentro de los eventos que no son cuantitativos, podemos encontrar que algunos notarios mienten cuando registran algún documento, esto no puede ser calculado porque no es cuantitativo, en el sentido de que si sólo sabemos que una vez mintieron; pero no sabemos cuántas veces más han mentido. Por ello, podemos asignar una probabilidad a este evento dependiendo de cómo lo interpretemos. Algunas personas creerán que se les debe asignar una probabilidad, de más de ½, otros creerán que debe ser de menos.

- Su probabilidad es fija y si ésta cambia lo hará con base en ejemplificaciones y no en pruebas que son fundamentales en los métodos científicos para confirmar hipótesis.
- Tiene el mismo problema que el bayesiano, a saber, determinar la probabilidad inicial o *a priori.* El método frecuencialista no considera este problema como propio de la probabilidad, por ello, no lo trata, no obstante éste afecta en los resultados que da en el momento de su aplicación. Cabe resaltar con respecto al último punto que el método bayesiano trata el problema de la probabilidad inicial y a lo largo de esta investigación se expondrá cómo es que trata de solucionar este problema, aunque no lo logre.

Ahora podemos decir que estas ventajas y desventajas de los dos métodos de probabilidad que vamos a destacar en esta tesis, serán tomadas como premisas que apoyen la tesis principal de esta investigación, donde se han destacado a las inferencias científicas, cuya estructura o forma es de causa-efecto pues, por lo general, es la que más se usa en los métodos científicos para confirmar hipótesis, en particular, en el método hipotético deductivo científico.

Así podemos decir que entre las razones que se darán para apoyar la hipótesis principal de este trabajo está que el método hipotético deductivo científico tiene principalmente dos problemas relacionados con la probabilidad, a saber, el problema de la evidencia y de la justificación del método. La teoría bayesiana tiene ventajas como, por ejemplo, pregunta e indaga sobre los criterios que se usan para tratar de justificar un método científico y cómo los criterios afectan en el cálculo de la probabilidad. En cambio, el método frecuencialista de la probabilidad no da razones para su justificación como método de probabilidad y mucho menos se pregunta por los criterios de justificación.

El método hipotético deductivo asigna probabilidades a las hipótesis, dada la evidencia que está compuesta por los eventos que nos interesan destacar en los experimentos que se realizan; de éstos hay algunos que se pueden repetir (tener una frecuencia de eventos) y otros no, es decir, hay algunos que no tienen frecuencia, pero sí tienen características cualitativas y pruebas anteriores (probabilidad *a priori*). En esta clase de experimentos el método bayesiano de la probabilidad tiene una gran utilidad, pues se apoya en la probabilidad *a priori,* para poder calcular la probabilidad de hipótesis en donde se realizan experimentos que no se pueden repetir. En cambio, el método frecuencialista de la probabilidad necesita una frecuencia (ejemplificaciones), es decir, un evento repetible para poder calcular la probabilidad. Por ello, podemos decir que no puede ser usada en experimentos que no se pueden repetir. Además, el método hipotético deductivo científico usa evidencia con características tanto cualitativas como cuantitativas para confirmar la hipótesis dada la evidencia; luego, sabemos que el método bayesiano puede tratar este tipo de evidencia, en cambio, el método frecuencialista no. Otra característica que tiene el método bayesiano es que se adecua a las características del método hipotético deductivo científico y a la evidencia con que trata, pues en el método hipotético deductivo científico las pruebas anteriores que se hayan realizado a una hipótesis, antes de aplicarle el método científico, dan como resultado una probabilidad previa no desdeñable a la hipótesis, esta probabilidad sólo es calculada por el método bayesiano y es relevante para el método hipotético deductivo, para asignar una probabilidad a la hipótesis en cuestión. Utilizando estas razones, podemos concluir que el método bayesiano es más ventajoso en comparación con el método frecuencialista cuando se aplica al método hipotético deductivo para confirmar hipótesis científicas. Para reforzar este argumento, a favor de la hipótesis principal de esta tesis, incluiré ejemplos en donde se muestre que el método bayesiano es más ventajoso que el método frecuencialista de la probabilidad.

A pesar de centrarnos sólo en el método hipotético deductivo científico, cabe destacar que vamos a contrastar éste método con el método de falsificación y el método de confirmación. Estos métodos también serán destacados para hacer énfasis sobre el uso del método frecuencialista o bayesiano de la probabilidad.

Vale la pena anunciar que en lo que sigue sólo nos referiremos al contexto de justificación de las teorías científicas y no a su contexto de descubrimiento. De los varios problemas que puede haber en los métodos científicos que usen probabilidad, sólo nos enfocaremos en el problema de su justificación, en el problema de la evidencia y en el manejo de ésta, para la confirmación de hipótesis en los métodos científicos en general. Esto lo haremos, porque estos problemas también son parecidos a los problemas de la teoría de probabilidad bayesiano y de la teoría frecuencialista, pues ambas tienen el problema de su justificación, es decir, que los varios conceptos que introducen no están definidos ampliamente y esto hace que cada una de sus teorías de la probabilidad sean consideradas como no eficientes, pues los conceptos que se utilizan son para calcular la probabilidad bajo su interpretación de probabilidad. En el caso de la teoría bayesiana, hay una interpretación subjetiva bayesiana, en cambio, la frecuencialista tiene una interpretación objetivista de la probabilidad.

Tampoco se discute sobre si el conocimiento es falible o no, pues puede haber varios puntos de vista sobre el conocimiento, uno de ellos es que se cree que puede cambiar, es decir, ser falible y este puede ser un tema controversial. Pero aquí sólo supondremos que el conocimiento es falible, es decir, que está en desarrollo y se está perfeccionando; no entraremos en discusiones filosóficas al respecto. Por ejemplo, cuando hable sobre el conocimiento y/o cite a Bertrand Russell, sólo me enfocaré en cómo afecta en el conocimiento que tenemos sobre los eventos que suceden en el cálculo de la probabilidad, es decir, para un científico que tiene una hipótesis *x,* un evento *y* puede ser evidencia relevante para apoyar su hipótesis científica, cuando se calcule su probabilidad dada la evidencia, sin embargo, otro científico puede discrepar y decir que no es evidencia relevante o ni siquiera considera que sea evidencia. Sólo destacaré lo que dice Russell con respecto a la probabilidad y el conocimiento. En el ejemplo anterior, hay varios factores y circunstancias que influyen en nuestro conocimiento y, por ende, en cómo es que podemos estar de acuerdo en que algo puede ser evidencia relevante o no. Así pues, sólo destacaré esta parte que se refiere al conocimiento, pues es relevante mencionar que el tipo de evidencia que usemos para calcular la probabilidad puede hacer que está cambie en una hipótesis dada la evidencia. Además, el tipo de evidencia que usemos depende de qué es lo que consideremos como evidencia y en específico la evidencia relevante; luego, el tipo de evidencia, que es considerada como evidencia, depende de nuestro conocimiento, de cómo factores y circunstancias afectan a nuestro conocimiento.

Tampoco se discutirán las distintas interpretaciones sobre la probabilidad, sólo mencionarán aquellas referentes a los métodos de probabilidad bayesiano y frecuencialista, pues detenernos en cada interpretación de la probabilidad no es relevante para el objetivo de esta tesis. Por ello, sólo nos concentraremos en las interpretaciones objetiva, cuantitativa, subjetiva y cualitativa de la probabilidad.

No vamos a discutir ni expondremos problemas sobre las fórmulas de la probabilidad o problemas matemáticos que se relacionen con el uso de los métodos bayesiano y frecuencialista de la probabilidad, pues hacerlo involucraría exponer su prueba matemática y pruebas sobre los conceptos que son usados. Esto no será explicado porque hay pruebas que son tan amplias que involucran recurrir a otras fórmulas matemáticas que tendríamos que explicar, lo que resultaría poco productivo para el propósito de esta investigación; aunque, sí se darán referencias de textos en donde se pueden consultar. Lo que si haremos, es exponer la parte formal indispensable de los métodos de la probabilidad bayesiano y frecuencialista, para entender las ventajas y desventajas de su forma formal, las cuales destacaremos en esta investigación.

No introduciremos problemas de argumentos inductivos o de lógica inductiva, sólo nos limitaremos a mencionar el problema de justificación de la inducción y sus efectos en los métodos científicos, esto se hará para ver cómo la probabilidad como método trata este problema. Esto con el fin de exponer la propuesta de justificación que proporciona la probabilidad para los argumentos inductivos, que por lo general son usados en los métodos científicos, en específico, el método hipotético deductivo para la confirmación de hipótesis científicas.

Esta investigación desarrollará sus capítulos de la siguiente manera:

En el capítulo 1 se expondrán los problemas clásicos de la probabilidad cuantitativa y cualitativa, esto se hará para que se note cómo, a lo largo de la historia, fue evolucionando la teoría subjetiva y objetiva de la probabilidad, hasta llegar a las teorías bayesiana y frecuencialista de la probabilidad. Para hacer esta exposición, primero se mencionará por qué surgieron estas dos vertientes de la probabilidad cualitativa y cuantitativa, que nos interesan destacar. Después se explicará la vertiente cuantitativa con: el cálculo aritmético de Galileo, la probabilidad clásica (Laplace) y la evolución de esta vertiente hasta el frecuencialismo. Luego se explicará la vertiente cualitativa con los siguientes temas: la objeción de Maistrov sobre el cálculo aritmético de Galileo, la probabilidad epistemológica y, por último, el bayesianismo. En este último tema, se explicarán varios conceptos que son necesarios para entender la regla de Bayes; la cual es fundamental explicar para entender el bayesianismo; a lo largo de esta explicación se darán varios ejemplos para apoyar el entendimiento de varios conceptos formales. Asimismo, se destacarán varios atributos que la regla de Bayes tiene y que son considerados como ventajas cuando es usada; los cuales serán una premisa para apoyar la hipótesis principal de esta investigación.

También se mencionarán los axiomas de la probabilidad, con el fin de mostrar que no cualquier teoría de la probabilidad puede ser considerada como rigurosa, pues si no satisface por lo menos los axiomas de la probabilidad no será considerada. Asimismo, se mencionarán los axiomas como apoyo para la explicación de los atributos formales que tiene tanto el método bayesiano como el frecuencialista. Es importante señalar estos atributos, porque serán un apoyo para las premisas que estén a favor del uso del método bayesiano o frecuencialista de la probabilidad.

En el capítulo 2 se expondrán y analizarán las dos posturas de la probabilidad, que también son conocidas como la teoría subjetiva y objetiva de la probabilidad. Se dará una explicación de las distintas formas de clasificar las interpretaciones filosóficas de la probabilidad y cómo las características que se destacan y son tomadas en cuenta para explicar tales teorías, son clasificadas bajo cierto grupo; esto se hará para destacar las ventajas y desventajas de asumir una teoría frecuencialista o bayesiana de la probabilidad cuando se usan para la confirmación de hipótesis. Después, se expondrá la teoría frecuencialista de la probabilidad y la teoría bayesiana de la probabilidad, esto se hará con el fin de exponer lo que las distingue entre ellas y con las demás interpretaciones de la probabilidad. En la interpretación filosófica de la probabilidad, ya sea la del método bayesiano o la del frecuencialista, son distintas las interpretaciones de los conceptos que hay en los métodos de la probabilidad, veremos con más detalle esta interpretación filosófica en los capítulos 3 y 4 de la tesis.

En el capítulo 3 después de haber analizado la parte formal filosófica de los métodos de probabilidad, nos ubicaremos en condiciones para abordar la discusión filosófica relacionada con la contrastación de hipótesis científicas, pues necesitamos los elementos formales dados en el capítulo 1, para poder exponer la aplicación del método bayesiano de la probabilidad en los métodos científicos para la confirmación de hipótesis. Se expondrán los problemas que hay tanto en los métodos científicos que usan probabilidad para confirmar hipótesis científicas, como en los métodos de probabilidad usados para tratar los problemas y en particular me referiré al método hipotético deductivo científico. Dichos problemas son la justificación tanto de los métodos científicos como de los métodos de probabilidad, el problema de la interpretación de la evidencia, la recolección de la evidencia y la selección de la evidencia relevante. Estos temas se expondrán para ver cómo afectan ellos mismos en los resultados de los métodos de la probabilidad y así ver cómo éstos a su vez afectan en la asignación de probabilidad en las varias hipótesis que se tengan sobre cierto evento *x*. También se expondrá el uso de la probabilidad en los métodos científicos en general, esto se hará para dejar claro cómo se usa la probabilidad en los métodos científicos.

En el capítulo 4 se contrastarán las ventajas y desventajas cuando se usan los métodos bayesianos bajo cierta interpretación de la probabilidad en los métodos científicos. Esto se hará para exponer las propuestas que se puedan dar para solucionar las desventajas; de esta forma se dará una razón para considerarlos más eficaces en su uso, para la confirmación de hipótesis científicas en los métodos científicos. En particular, se destacará, como ya se mencionó, el método hipotético deductivo y los problemas que involucran su uso, esto se hará para exponer las posibles soluciones que puedan aportar éstos, ya sea el método bayesiano o el método de la frecuencialista de la probabilidad. Se darán ejemplos generales del uso de los métodos bayesiano y frecuencialista de la probabilidad en los métodos científicos y, en específico, del método hipotético deductivo científico. Todo lo que se verá en este capítulo se hará para mostrar en qué aspectos es mejor usar uno de los dos métodos de probabilidad ya mencionados.

En el capítulo 5 se confrontarán las razones que hemos expuesto a lo largo de todo este texto para apoyar la hipótesis principal de esta investigación, a saber, que es que el método bayesiano de la probabilidad tiene más utilidades que el método frecuencialista de la probabilidad en los métodos científicos, en ciertas circunstancias, para confirmar hipótesis y específicamente en los métodos científicos que tienen la estructura causa-efecto, en particular el método hipotético deductivo.

Ahora, es fácil advertir qué importancia tiene abordar y desarrollar la tesis principal de este texto en el ámbito científico, específicamente en donde se usan los métodos científicos para confirmar hipótesis, pues hacemos explícitas posibles soluciones a problemas que tiene el método científico para corroborar hipótesis; esto lo hacemos cuando exponemos las ventajas y desventajas que tiene el método bayesiano y el método frecuencialista y cuando son utilizados en los métodos científicos, pues al mostrar estas ventajas podemos contrastarlas con las posibles soluciones que pueden darse a los problemas que tiene el método científico en general; cuando hacemos esto último concluimos razones para argumentar que el método bayesiano tiene más utilidad que el frecuencialista de la probabilidad en ciertas circunstancias y problemas que tiene el método científico, específicamente los métodos que tiene la estructura causa-efecto, en particular, el método hipotético deductivo.

De lo anteriormente dicho, podemos notar que cuando concluimos estas razones, para argumentar a favor de un método probabilístico, proporcionamos al científico razones rigurosas que justifiquen la aplicación de un método probabilístico a un método científico para confirmar hipótesis. Así pues, cuando el científico tiene buenas razones para aplicar cierto método de la probabilidad, puede obtener una probabilidad más acotada y mejor fundamentada o justificada desde el cálculo de la probabilidad de una hipótesis, dada la evidencia. De esta manera, el científico tiene menos errores y consecuencias imprevistas de la aplicación de una hipótesis, cuando ésta es confirmada, pues obtuvo una probabilidad más acotada.

I. Métodos o enfoques de la probabilidad

Introducción

En este capítulo se expondrá de forma general y breve la historia de la probabilidad, esto se hará con el fin de mostrar cómo desde los inicios del estudio de la probabilidad se comienzan a distinguir dos tipos de probabilidad: la cualitativa y la cuantitativa. Después se mencionarán la probabilidad clásica y la probabilidad axiomática, pues con base en éstas podremos entender las características que distinguen al método bayesiano de la probabilidad del método frecuencialista, estos métodos también serán expuestos de una manera formal.

Históricamente ha habido dos vertientes en el estudio de la probabilidad, a saber, la vertiente cuantitativa también conocida como objetiva y la vertiente cualitativa también conocida como subjetiva. A lo largo de la historia de la probabilidad, diferentes teorías la han entendido de maneras diversas, dentro de esa diversidad las dos vertientes mencionadas son las más destacadas.

1.1 La vertiente cuantitativa de la probabilidad

1.1.1 Galileo y el cálculo aritmético de la probabilidad

La probabilidad es un método que ha sido empleado en lo más cotidiano de la vida humana, como para buscar ventajas en los juegos de azar[6]. Pero también en los estudios con un carácter científico como en la astronomía que es una ciencia muy antigua o como en inteligencia artificial; en este texto sólo nos limitaremos a tratar la probabilidad dentro de estudios con carácter científico. Ahora bien, cuando los métodos probabilísticos son

6 Hacking afirma que el origen de la probabilidad se encuentra en los fenómenos aleatorios, juegos de azar y en la astrología. Los primeros pasos de la probabilidad fueron registrados a partir del siglo XVI en un primer texto *juegos de azar*. [Hacking 1995, p.56]

aplicados a tipos de evidencia con distintas características, se originan distintos resultados y, por ende, distintas opiniones de los científicos sobre el por qué los resultados son distintos y cuál es el más acorde a la realidad. De estos distintos resultados originaron distintas posturas sobre la probabilidad, en este primer subcapítulo me interesa destacar dos posturas sobre la probabilidad (la subjetiva y la objetiva), pues exponen las características con las que se distinguen e identifican la teoría bayesiana de la probabilidad y la teoría frecuencialista de la probabilidad; estas teorías de la probabilidad deben ser expuestas, para que de esta manera podamos ver cómo afectan en los resultados de la aplicabilidad de los métodos de probabilidad bayesiano y frecuencialista cuando son utilizados en los métodos científicos para la corroboración o refutación de hipótesis. Las teorías de la probabilidad bayesiana y frecuencialista se distinguen por tratar con distintos tipos de evidencia, a saber, evidencia con características cualitativas y evidencia con características cuantitativas[7]; esta distinción ha sido tan característica de las distintas posturas de la probabilidad que se les puede reconocer como probabilidad cuantitativa[8] y probabilidad cualitativa, respectivamente.

La probabilidad cuantitativa puede ser usada para saber con razones aritméticas qué eventos son más probables; podemos citar por caso los ejemplos expuestos por Hacking sobre los juegos de azar, uno de estos expone claramente cómo por medio de cálculos

---

[7] A lo largo de la historia de la probabilidad se pueden distinguir dos términos que señalan a dos tipos de posturas de la probabilidad, a saber, la postura objetiva de la probabilidad y la subjetiva de la probabilidad, ambas serán enfatizadas y desarrolladas a lo largo de este texto, pues serán importantes para ubicar las posturas filosóficas de los dos métodos de probabilidades que se van a desarrollar. La probabilidad cualitativa se distingue por ser menos formal que la probabilidad cuantitativa, más adelante veremos con más detalle tal distinción.

[8] Hacking usa el término ‘probabilidad cuantitativa’ para referirse a cualquier teoría de la probabilidad que tenga la característica de tratar sólo con eventos enumerativos, aquella probabilidad que maneja números para saber qué casos o eventos son más probables que otros; este tipo de probabilidad hace referencia a la teoría frecuencialista de la probabilidad que a su vez es clasificada como una teoría objetiva de la probabilidad. Coincido con su clasificación.

podemos saber qué números son más probables en una tirada de dados y por qué. De acuerdo con Hacking, antes de Galileo se creía que sólo por medio de la observación exhaustiva se sabía qué números era más probable que salieran en una tirada de dados; esto fue así, hasta que Galileo probó por medio de razones aritméticas que no era necesaria la observación exhaustiva para saberlo[9].

1.1.2 Laplace y la probabilidad clásica

De esta forma la probabilidad cuantitativa se inclinó por explicar la probabilidad por medio de razones aritméticas. Después, el filósofo y matemático Laplace expuso un método probabilístico que da razones de manera rigurosa para explicar por qué un fenómeno aleatorio ocurre con más frecuencia que otro.

A continuación se expondrá 'la probabilidad clásica' que es considerada como una de las primeras interpretaciones de la probabilidad y fue propuesta por Laplace. Su definición de la probabilidad está compuesta de principios que son esenciales para entender la postura que expone en su texto: *Ensayo filosófico sobre las probabilidades.* Los términos que se verán son importantes para entender cuál fue la propuesta del método frecuencialista laplaciano.

Laplace. Equiprobabilidad y Equiposibilidad

Laplace decía que la teoría de la probabilidad es sólo sentido común expresado con números.[10] Uno de sus principios hace referencia a las opciones de casos y la elección que

---

[9] El ejemplo de Galileo se encuentra en [Hacking 1995, pp.70-71].

[10] En efecto, como vimos en el ejemplo de Galileo, ya en los juegos de azar, los jugadores intuían qué números sucedían más, pero no sabían, a ciencia cierta, por qué cierto evento aleatorio ocurría con más frecuencia que otro; sólo sabían por experiencia colectiva y/o personal que el evento sucedería con más

hacemos, de uno de ellos, sin tener razones para elegir uno sobre los otros. Este principio es el 'principio de indiferencia', así nombrado porque somos indiferentes al elegir algún caso de entre los que tenemos. En otras palabras, "Este principio establece que dos posibilidades son igualmente probables, pues no hay razón para preferir una sobre otra."[11]

El principio de la indiferencia sugiere que si no tenemos razones para elegir un evento sobre los demás, podemos considerarlos como igualmente probables, es decir, como eventos equiposibles. De este modo Laplace define los casos que son iguales como 'equiposibles' o 'equiprobables'; bajo este supuesto establece la definición de probabilidad clásica como la razón del número de casos favorables sobre el de todos los casos posibles. En palabras de M. A. García:

> La elección al azar se refiere a experimentos aleatorios en los cuales se dispone de una colección de objetos, los cuales pueden ser bolas, cajas, tarjetas, personas etc., y se define el experimento aleatorio precisamente como la elección de uno o varios objetos de la colección. El término 'elegir al azar' que se utiliza en esos casos se entiende en el sentido de que, al elegir los objetos, no se tiene preferencia por la elección de uno de ellos sobre la elección de otro. [García 2005, p. 54]

Según esta definición, la probabilidad de un evento *A* se calcula de la siguiente manera:

$$P(A) = \frac{\#\ \text{de eventos elementales que producen la ocurrencia de } A}{\#\ \text{total de eventos elementales}}$$

[García 2005, p. 54]

Otra manera de representarlo, en notación moderna:

---

frecuencia. El sentido común era la única razón por la que se apostaba a ciertos números. De ahí el dicho de Laplace: "En el fondo la teoría de probabilidades es sólo sentido común expresado con números."[Laplace en Andrew 1994, p.3] Uno de los principios de Laplace se refiere a la ignorancia humana, de hecho en su teoría de la probabilidad se enfoca en la ignorancia del hombre. Su principio afirma que nuestro conocimiento es limitado, es decir, a veces no tenemos el conocimiento para saber cuándo va a ocurrir cierto fenómeno, pero podemos saber que ocurrirá. Dice Laplace: "Todo evento es predecible hasta el grado en que la ignorancia humana lo permita." [Laplace en Andrew 1994, p.3] De ello se sigue que la probabilidad es relativa, en parte a nuestra ignorancia y en parte a nuestro conocimiento.

[11] [Salmon 1985, p. 65]

> Si un espacio muestral Ω consiste de n eventos elementales equiposibles y ε_i es un evento contenido en Ω, $\varepsilon_i \subset \Omega$, con n_ε resultados favorables a ε_i, entonces la probabilidad de ε_i está dada por la fórmula.
>
> $[\, P(\varepsilon_i)= n_\varepsilon / n]$ [Hernández 2007, p. 42]

La probabilidad, así definida, cumple con todas las propiedades que Laplace llamó 'principios' en su texto *Ensayo filosófico de las probabilidades*[12]. Los principios son: (1) nuestro conocimiento es limitado; (2) todo evento es predecible hasta el grado en que la ignorancia humana lo permita; (3) todos los eventos son gobernados por causas necesarias.

Laplace creía que todo era en principio susceptible de ser medido, en contraste sus críticos acotarían el campo de acción de la probabilidad únicamente a aquellos casos que formaran una secuencia de observaciones repetibles y aleatorias a la que llamaron "colectivo". Con esto la definición de probabilidad cambió en 1928 por von Mises a:

> Sea A un atributo arbitrario de un colectivo C obtenido m veces en n intentos; la probabilidad de A dado C es:
>
> $$P(A|C)= \lim_{n \to \infty} \frac{m(A)}{n} \quad (n \text{ tiende a infinito})$$[13]

---

[12] La probabilidad clásica laplaciana a sido criticada por lo menos en dos de sus facetas:

a) La idea de 'equiposibilidad' o 'equiprobabilidad'.
b) La idea de suponer que todos los casos son posibles

El primer problema es que no queda del todo claro si se distingue equiposibilidad de equiprobabilidad y cómo. El segundo problema es el de no considerar la frecuencia con que ocurre un evento en la práctica, pues la probabilidad clásica no considera los experimentos que se realizan sobre un evento del que se quiere calcular su probabilidad; por ejemplo, en el juego de los volados (el lanzamiento de una moneda) para la probabilidad clásica, la probabilidad de que salga águila en el siguiente lanzamiento es de ½, pues la moneda tiene dos posibles resultados (águila o sol); en cambio para el frecuencialista (uno de sus críticos) la probabilidad de que en el siguiente lanzamiento salga águila se calcula con mayor precisión no suponiendo de antemano que es igualmente probable que caiga águila que sol, sino por medio de la observación de resultados de sucesivos lanzamientos de la moneda. Así pues, surgieron críticos que empezaron a decir que la probabilidad no se trata de la medida de 'todos los casos equiposibles', sino más bien de la frecuencia hacia la cual tendería una secuencia infinita, esto lo dice la teoría frecuencialista de la probabilidad.

[13] No se puede definir el límite que tiende a infinito, según M.A. García; si es así, encontramos una deficiencia en la nueva definición. Ahora bien, no se puede definir porque el cálculo tiende a una probabilidad exacta de un evento A pero nunca llega a la probabilidad exacta sólo se acerca (cuando digo exacta me refiero a la postura de los frecuencialistas, pues ellos creen que un evento *A* sólo le corresponde una probabilidad).

Esta definición fue la fórmula que reemplazó la definición de Laplace y es la que caracteriza a la teoría frecuencialista de la probabilidad. De esta forma, podemos ver que la probabilidad clásica dio paso a la probabilidad frecuencialista, pero las dos siguen bajo la misma vertiente de la probabilidad, la vertiente cuantitativa.

Ahora bien, lo que está calculando el frecuencialista no es lo mismo que calcula la probabilidad clásica, porque a la probabilidad frecuencialista le interesa la probabilidad conforme a los hechos ocurridos. En la fórmula anterior se quiere calcular la probabilidad exacta de un evento A, pero ésto no es posible porque lo que se calcula es un límite en el que n tiende a infinito, es decir, con la fórmula de arriba los frecuencialista sólo pueden acercarse a la probabilidad exacta pero nunca llegan.

Sin embargo tratan de solucionar este problema diciendo que el resultado obtenido por medio de la fórmula se aproxima a la probabilidad exacta de A, es decir, tiende a $P(A)$.

1.1.3 El frecuencialismo

Una característica esencial del método frecuencialista de la probabilidad es la definición de "colección" o "colectivo", pues este método de la probabilidad sólo trabaja con colecciones de eventos, es decir, no se puede aplicar el método frecuencialista hasta que se ha determinado una colección. *Colectivo* o *colección* se define en probabilidad como una secuencia de eventos uniformes o procesos que difieren entre sí en ciertos atributos observables; además las colecciones deben satisfacer el requisito de ser totalmente aleatorias. Por ejemplo: una colección estaría conformada por las personas que sacan su credencial de elector al año, o por todas las tiradas de dados que se realizan en un juego donde el atributo de cada evento es el número que se obtuvo. Por el contrario, no constituyen colecciones cuántas veces un individuo en específico se cae de las escaleras al año, o algo que parece imposible: cuántas veces nace. Es decir, cuando pensamos en colecciones no debemos pensar en individuos, sino más bien en una cierta clase considerada como un todo.

Una característica propia del método frecuencialista es el concepto de *distribución binomial*. Dicho concepto fue expuesto por Jacob Bernoulli, quien atribuía ciertas características a un experimento aleatorio, estas características son: que en una frecuencia de experimentos sólo se tiene éxito o fracaso, es decir, que ocurra o que no ocurra lo que se espera, es constante. Un experimento sólo es binomial si hay un número fijo de experimentos independientes. Otro concepto particular del método frecuencialista es el de *frecuencia relativa,* definido como el cociente m/n, en donde m es el número de veces en que aparece el atributo que queremos reconocer en el evento que ocurre n veces. El valor que resulte de la frecuencia relativa es la probabilidad de que cierto atributo aparezca en una colección de eventos; hemos de notar que la experiencia nos muestra que en un colectivo de eventos que representan las frecuencias relativas de ciertos atributos se vuelven más predecibles o estables a medida que el número de observaciones aumenta.

Ahora bien, el conjunto de todos los posibles resultados de un experimento se llama *espacio muestral*; un evento se define como un subconjunto particular del espacio muestral; para cualquier evento sólo puede presentarse una de dos posibilidades: que ocurra o que no ocurra. En resumen, la frecuencia relativa de ocurrencia de un evento, en un cierto número de repeticiones del experimento, es una medida de la probabilidad de tal evento, es decir, a medida que un experimento aleatorio se repite un mayor número de veces, los posibles resultados tienden a presentarse cada vez más cerca, lo cual indica que la frecuencia de aparición de cada resultado tiende a estabilizarse[14].

Debemos mencionar que en la interpretación de la probabilidad frecuencialista lo probable son los eventos considerados como subconjuntos de colecciones. En este enfoque, la frecuencia relativa está relacionada con el posible resultado de un evento relativo a un experimento en cuestión y es una fracción que resulta de dividir el número de ocasiones en que un evento se ve favorecido, al realizar una serie de experimentos, entre el total de experimentos de los que consta la serie. La definición frecuencialista de las probabilidades da lugar al siguiente principio:

> A esta propiedad de regularidad de la frecuencia relativa con que ocurre cada uno de los posibles resultados de un experimento aleatorio repetible o, en general, cada evento relativo al experimento, lo llamaremos principio de regularidad de las frecuencias. [García 2005, p.38]

El principio de regularidad de las frecuencias expresa que la frecuencia relativa con la que ocurre cada posible resultado en un experimento aleatorio converge hacia una constante cuando la serie de experimentos es relativamente grande. Esta idea ha sido representada por un límite que tiende a una constante como antes se ha señalado.

Desde la perspectiva matemática, la probabilidad frecuencialista se puede definir de la siguiente forma:

---

[14] Un ejemplo de ello sería: en cien juegos de volados, no sabemos exactamente cuántas veces una moneda caerá águila ni cuántas veces caerá sol, hasta que realicemos el experimento. Suponiendo que las veces que cayó águila fueron 47 y las veces que cayó sol fueron 53; si dividimos 47 entre 100 y 53 entre 100 obtendremos 0.47 y 0.53, estos resultados se acercan a 0.5. Ahora bien, si queremos acercarnos más a 0.5, debemos repetir el experimento mucho más veces. Si lanzáramos la moneda 100000 veces y obtuviéramos como resultado que cayeron 50008 águilas y 49992 soles, entonces tendríamos de resultado un número que se acerca más a 0.5, pues si dividimos 50008 entre 100000, obtendríamos 0.50008, éste es un número que se acerca más a 0.5. Podemos decir que entre más experimentos aleatorios se realicen, el resultado tenderá a 0.5.

> Si A es un atributo arbitrario asociado con una colección particular; si Ω es el espacio muestral atribuido al colectivo, entonces $A \subset \Omega$. Supóngase esto en los primeros n miembros de la colección A ocurridos $m(A)$ veces, sabiendo esto, la frecuencia relativa es $m(A)/n$ y consigue un valor cada vez más cerca de un valor fijo. [Gillies 2000, p. 92]

En otras palabras:

> Este enfoque asume que si un experimento se repite n veces, de las cuales n_A ocurre el evento A, entonces la probabilidad se define como el límite de la frecuencia relativa n_A/n lo cual quiere decir que, [en este ejemplo]. [Hernández 2007, p.43]

$$\frac{n_A}{n} \rightarrow P(A)$$

El inconveniente de esta interpretación de la probabilidad es que para calcularla se tiene que realizar un experimento aleatorio una cierta cantidad de veces, lo que restringe su uso a experimentos que pueden ser repetidos; esta limitación puede ser vista como una desventaja en su uso, específicamente, en los métodos científicos, más adelante veremos por qué es una limitación. Dicho de manera concisa, el problema es que el cálculo de la probabilidad frecuencialista depende de poder hacer n observaciones (donde n debe ser relativamente grande), pero realizar dichas observaciones puede estar restringido por circunstancias y factores ya sean económicos o naturales. Por ejemplo, cuando se realiza en experimento que consta de golpear un átomo contra otro, el problema es que el experimento es tan costoso que no se puede repetir varias veces.

Hasta aquí podemos notar cómo ha evolucionado la vertiente cuantitativa de la probabilidad, desde Galileo con su cálculo aritmético que dio paso a la probabilidad clásica y que ésta a su vez dio paso al método frecuencialista de la probabilidad.

1.2 La vertiente cualitativa de la probabilidad

1.2.1 La objeción de Maistrov

La vertiente cuantitativa de la probabilidad, desde el cálculo aritmético de Galileo hasta el método frecuencialista, ha sido objeto de varias objeciones por parte de proponentes de la vertiente cualitativa. En el caso del cálculo aritmético de Galileo, Maistrov[15] expresa una limitación basada en que hechos empíricos que no pueden ser contados pueden afectar el cálculo de la probabilidad. Por ejemplo, cuando en una tirada de dados calculamos la probabilidad de que salgan ciertos números, esta probabilidad puede ser afectada por factores empíricos, en este caso puede ser que los dados estén trucados.

Este es uno de los problemas más importantes de la probabilidad clásica, especialmente si deseamos manejar cálculos aritméticos para solucionar situaciones prácticas; por ejemplo, al querer saber cuál sería la repartición equitativa de un premio entre dos equipos que estaban jugando pero no pudieron terminar por cuestiones azarosas. En un caso así, los hechos empíricos pueden afectar el resultado aritmético obtenido al aplicar un método de la probabilidad. Pero lo interesante es saber que aunque los eventos no enumerativos pueden ser eventos que no se repiten, sin embargo, afectan en las decisiones de cómo repartir el premio en el ejemplo anterior.

1.2.2 De Pascal a la probabilidad epistemológica

Pascal sugirió una aritmética de la probabilidad que no estuviera basada en juegos de azar y que fuera más general en su aplicación, es decir, que también fuera práctica en la toma de decisiones en la vida común humana o en las decisiones que se toman en la ciencia. Este tipo de probabilidad con características epistemológicas fue llamada *probabilidad epistemológica* o *probabilidad cualitativa*. La probabilidad epistemológica se distingue de la probabilidad cuantitativa, porque en ella se puede comparar el grado por el cual la evidencia garantiza diversas proposiciones sin recurrir a los números, en cambio, la

---

[15] [Maistrov en Hacking 1995, p.71]

probabilidad cuantitativa necesita recurrir a los números para garantizar diversas proposiciones.

Así pues, una expresión de probabilidad con características epistemológicas[16] fue utilizada para denotar algo medible desde la sugerencia de Pascal[17]. Sin embargo, la probabilidad epistemológica se desarrolló y se distinguió de la probabilidad cualitativa, porque es considerada como: "aquella que compara el grado en el cual la evidencia garantiza diversas proposiciones sin recurrir a los números."[18]

El uso de la probabilidad epistemológica y la evidencia

El manejo de la evidencia en la probabilidad epistemológica para decidir si una opción de las que podemos tener es más útil que otra es considerada como un tema importante en la filosofía de la ciencia y en la vida real. Así pues, podemos decir que la probabilidad epistemológica se refiere a la vertiente cualitativa de la probabilidad. Un ejemplo donde podemos ver claramente la distinción de la probabilidad cualitativa y cuantitativa es el siguiente.

Cuando se tienen varias observaciones que sirven como evidencia para una cierta hipótesis pueden surgir problemas, porque puede suceder que la evidencia *x* esté apoyando la hipótesis, pero la evidencia *y* la rechaza; un ejemplo de ello es el expuesto por Ian Hacking:

---

16 Asignar probabilidades es visto como una forma de evaluar epistemológicamente y, en este texto nos interesa saber, que por medio de la evaluación de la probabilidad, saber qué hipótesis científica corroborar o invalidar. Sin embargo, como hicimos notar en la introducción de este texto, la utilidad de la probabilidad depende de cómo se aplique y a qué tipo de evidencia, además de con qué método de la probabilidad.

17 Russell y su concepto de grados de credibilidad también tienen características epistemológicas sobre la probabilidad y expresa una postura subjetiva de la probabilidad, la cual, veremos más adelante, es tan útil como la probabilidad matemática, es decir, la probabilidad axiomática que refiere Russell.

18 [Hacking 1995, p.6]

> Considérese la pregunta sobre si un contrato del que dos notarios dan fe ha sido postdatado. Como es cierto que 999 de 1000 contratos debidamente notariados han sido fechados correctamente, entonces, "si no conocemos ninguna otra particularidad sobre el contrato", deberíamos creer que la fecha es honesta. "Es incomparablemente más probable que el contrato ante mí sea uno de los 999 y no el único de los 1000 que está postdatado." No obstante si nos enteramos que los notarios son inescrupulosos, el documento se convierte en menos creíble. Si nos enteramos que al momento de la fecha del contrato la parte que debía prestar 20.000 libras sólo tenía 100 libras, entonces "creería que hay algo falso en el contrato". [Hacking 1995, p.101]

En el ejemplo anterior, Hacking trata de manifestar que tenemos una probabilidad cuantitativa, pero los siguientes datos del ejemplo tienen una forma cualitativa y estos datos afectan la probabilidad cuantitativa de una forma cualitativa; la afectan reduciéndola hasta eliminarla. Este es el problema de cuando tenemos varias observaciones que sirven como evidencia y unas apoyan la hipótesis, mientras otras la contradicen. Ahora bien, con este ejemplo nos explica que podemos tener instancias en donde la probabilidad cuantitativa debilite la hipótesis de una probabilidad cualitativa o viceversa.

Por consiguiente, podemos notar en el ejemplo anterior que para tomar decisiones que implican consecuencias en nuestras acciones sería mejor considerar una variedad de evidencias. Ahora bien, creo que debemos pensar los problemas de la probabilidad como originados en parte de datos empíricos y en parte de datos aritméticos, pues la probabilidad es un cálculo formal que se aplica a eventos empíricos observables de la vida real. Por ello, para obtener resultados acordes a la realidad y útiles necesitamos considerar tanto el cálculo como a lo que se le aplica. Esto es importante en la filosofía de la ciencia, pues cuando se quiere calcular la probabilidad de una hipótesis científica dada la evidencia observada, el cálculo de probabilidad que realicemos podría dar distintos resultados si usamos distinta evidencia y, si no consideramos esto en la probabilidad podríamos pensar que el método de la probabilidad que usamos para calcular la probabilidad tiene un problema en su fórmula y, por ello, es ineficaz.

Sea cual sea el uso de la probabilidad, ya sea para problemas que impliquen números o para elecciones de la vida cotidiana o para la corroboración de hipótesis científicas, no podemos negar que aclara las hipótesis que tenemos y que podemos elegir una de ellas; expone de una manera formal la evidencia que nos permite tener un criterio para elegir la mejor opción. Además, la probabilidad tiene varias teorías que algunas tratan eventos cuantitativos, otras tratan eventos cualitativos y muy pocas teorías tratan tanto eventos cuantitativos como cualitativos. Por lo general, los métodos de la probabilidad reemplazan la simple confusión y la observación exhaustiva que nos enseña que tal hecho es más plausible que otro, como en el ejemplo de Galileo que nos da razones aritméticas para creer en la probabilidad que se dé en un evento azaroso.

1.2.3 El bayesianismo

En la presente sección se introducirá la aproximación bayesiana a la probabilidad, que se clasifica bajo la vertiente cualitativa de la probabilidad porque es una teoría que incluye características subjetivas. En particular vamos a mencionar la regla de Bayes y algunos conceptos o reglas que la componen. Por otro lado, se apelará a estas reglas y conceptos con miras a argumentar que el método bayesiano es más eficaz que el método frecuencialista en los métodos científicos para la corroboración de hipótesis científicas en ciertas circunstancias y factores que explicaremos más adelante.

La regla de Bayes se caracteriza por calcular la probabilidad de un evento dada cierta información y ha sido muy útil – tanto, que ha sido aceptada por la teoría de la probabilidad objetiva como una regla eficaz. Para exponer la regla de Bayes se deben explicar varias reglas y nociones que influyen en ella, como la probabilidad condicional, la independencia estocástica y la regla de la probabilidad total. Después expondré las características que distinguen a la regla de Bayes de otras formas de calcular probabilidades y su interpretación de la probabilidad. Por último se mencionará la probabilidad inversa, el problema de Bayes y la solución que le da Bayes.

Probabilidad condicional

La probabilidad condicional se refiere a la tripleta (Ω, a, P) donde Ω es el espacio muestral, a es el evento que se quiere destacar y encontrar su probabilidad y P es la función de probabilidad que cumple con las siguientes clausulas:

i) $P\colon a \rightarrow \mathbf{R}$ (es decir, P asigna a a un número real)
ii) $P(a) \geq 0$
iii) $P(\Omega)=1$
iv) P es σ-aditiva.

La última propiedad mencionada, que puede considerarse como fundamental para la función de probabilidad, trata de descomponer un evento a en eventos más simples (mutuamente excluyentes y la unión de todos ellos es a); la probabilidad del evento a es la suma de las probabilidades de los eventos más simples y el evento a se descompone en eventos elementales para que sea más fácil encontrar su probabilidad; como dice M.A. García.:

> De manera general, el método para asignar probabilidades a los eventos relativos a cualquier experimento aleatorio va de lo simple a lo complejo: primero se encuentra la probabilidad de una clase particular de eventos y, a partir de ahí, utilizando las propiedades de la función de probabilidad, se extiende ésta a una clase más amplia de eventos y después a familias cada vez más extensas. [García 2005, p.52]

La probabilidad condicional resuelve problemas en experimentos aleatorios que están compuestos de varias partes y que se realizan de forma sucesiva como, por ejemplo, el caso de la urna que a continuación veremos. Sabido esto, se inicia calculando las probabilidades de los eventos que suceden en cada parte del experimento y, como son sucesivos, el primero antecede al segundo y, a consecuencia de ello, lo que haya ocurrido en el primero afectará al segundo[19]; de esta manera, se condiciona la probabilidad a lo que ocurra en la primera parte del experimento, por ello, se le nombra regla de la probabilidad condicional. Por ejemplo:

Una urna contiene 2 bolas rojas enumeradas y 3 bolas blancas enumeradas; el experimento aleatorio a realizar está compuesto por dos partes: en la primera se extrae al azar una bola de la urna sin reemplazo; en la segunda parte del experimento se extrae al azar otra bola de la urna. Ahora bien, el espacio muestral del experimento contiene los posibles resultados del experimento aleatorio, es decir, según el experimento, puede haber sido cualquiera de las siguientes 20 combinaciones, donde Bb= 'bola blanca' y Br= 'bola roja': $\{(Bb_1, Bb_2)$, (Bb_2, Bb_1), (Bb_1, Bb_3), (Bb_3, Bb_1), (Bb_3, Bb_2), (Bb_2, Bb_3), (Br_1, Br_2), (Br_1, Bb_1), (Bb_1, Br_1), (Br_1, Bb_2), (Bb_2, Br_1), (Br_1, Bb_3), (Bb_3, Br_1), (Br_2, Br_1), (Br_2, Bb_1), (Bb_1, Br_2), (Br_2, Bb_2), (Bb_2, Br_2), (Br_2, Bb_3), $(Bb_3, Br_2)\}$.

Ahora bien, si en la primera parte del experimento se obtiene bola roja, con este dato los posibles resultados se reducen a los eventos en que se obtiene bola roja en la primera extracción: $\{(Br_1, Br_2)$, (Br_1, Bb_1), (Br_1, Bb_2), (Br_1, Bb_3), (Br_2, Br_1), (Br_2, Bb_1), (Br_2, Bb_2), $(Br_2, Bb_3)\}$; además, la probabilidad de sacar una bola de cualquier color cambia

[19] Aunque hay casos donde los eventos son independientes, donde la ocurrencia o no ocurrencia del primero no afecta al segundo. Esta idea se desarrollará más adelante.

en el segundo experimento, pues el experimento es sin reemplazo y todos los posibles resultados tienen distinta probabilidad[20].

Esto es, si simbolizamos los eventos como:

A: en la primera extracción se obtiene bola roja.

B: en la segunda extracción se obtiene la bola B_i.

Donde, definimos la probabilidad de la intersección de *A* y *B* como:

$$P\ (A \cap B) = P(A)\ P(B)$$

La probabilidad de sacar una bola ya sea roja o blanca en la segunda extracción, dado que en la primera extracción se obtuvo roja es: la probabilidad de *B* dado *A* que es igual a la probabilidad de la intersección de *A* y *B* entre la probabilidad de *A* y de ello se obtiene:

$$P(B|A) = \tfrac{1}{4}$$

La definición de probabilidad condicional mencionada por M.A. García es:

> Sean *A* y *B* dos eventos y supongamos $P(A) > 0$, se define la probabilidad condicional de *B*, dada la ocurrencia de *A*, $P(B|A)$, mediante la fórmula:
>
> $$\boxed{P(B|A) = \frac{P(A \cap B)}{P(A)}}$$
>
> [García 2005, p.6]

Con la definición de probabilidad condicional podemos representar los experimentos como (B_i, B_j), donde $i \neq j$ y B_i representa el resultado de la primera parte del experimento, y B_j la segunda parte. Hemos de resaltar que antes de realizar los experimentos la probabilidad de sacar una bola es equiprobable[21], de este modo la probabilidad del evento *B* depende de la

---

[20] Cuando se dice que es sin reemplazo se refiere que el espacio muestral que se tiene al principio no es el mismo cuando se realiza el segundo experimento aleatorio, pues, en este caso, se sacó la bola y ya no se regresa a la urna y esto afecta al espacio muestral que se tenía al principio.

[21] Cuando se utiliza el término 'equiprobable' se refiere a que todos los eventos tienen igual probabilidad.

ocurrencia del evento A. Pero puede haber casos en que la ocurrencia de un evento que antecede a otro no afecte al segundo evento; en esos casos los eventos que componen el experimento son independientes, es decir, que suceda el evento B no depende de que no suceda o suceda el evento A.

Ahora bien, si queremos calcular la probabilidad de sacar la bola blanca: Bb_2, dado que ya sabemos todos los datos del experimento, se calcula su probabilidad de la siguiente manera:

$$P(Bb_2 \mid A) = \frac{P(A \cap Bb_2)}{P(A)}$$

$$P(Bb_2 \mid A) = \frac{P(A) \cap P(Bb_2)}{P(A)}$$

$$P(Bb_2 \mid A) = \frac{(8/20 \cdot 2/8)}{(8/20)}$$

$$= \frac{16/160}{8/20}$$

$$= \frac{320}{1280} = \frac{1}{40}$$

Dado que los eventos son dependientes, para calcular la probabilidad de la intersección de dos eventos dependientes se hace con la regla del producto.

Regla del producto

La definición de probabilidad condicional también nos permite definir la probabilidad de una intersección:

$$\boxed{P(A \cap B) = P(B \mid A)\, P(A)}$$

Esta definición es conocida como la regla del producto, lo interesante de esta regla es que puede utilizarse para calcular la probabilidad de una intersección de eventos en donde la probabilidad condicional de la fórmula se puede obtener de forma directa[22].

Regla de la suma

La regla de la suma hace referencia a dos eventos que son mutuamente excluyentes; la probabilidad de la unión de estos es la suma de las probabilidades menos la probabilidad de la intersección de los dos eventos, es decir:

Si A y B son eventos mutuamente excluyentes (Se dice que dos eventos A y B son mutuamente excluyentes, si es imposible que ocurran ambos, $A \cap B = \phi$, en cualquier parte del experimento aleatorio en el que suceden o al que son relativos), entonces la probabilidad de ellos es:

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Donde $P(A \cap B) = 0$

Independencia estocástica

[22] Por ejemplo: Una urna contiene 15 bolas rojas y 10 bolas azules; se realiza un experimento aleatorio el cual consiste en extraer al azar 10 bolas de la urna; se quiere calcular la probabilidad de que la primera bola azul se obtenga en la quinta extracción.

Queremos calcular la probabilidad de sacar en las primeras cuatro extracciones del experimento puras bolas rojas y en la quinta extracción una bola azul. Si simbolizamos los eventos como S_k= 'se obtiene la bola azul en la k-ésima extracción' y R_k= 'se obtiene la bola roja en la k-ésima extracción', entonces la probabilidad de que se dé en las primeras cuatro extracciones bola roja y en la quinta extracción se dé una bola azul es igual a:

$P(R_1 \cap R_2 \cap R_3 \cap R_4 \cap S_5)$

$= P(S_5 \mid R_1 \cap R_2 \cap R_3 \cap R_4)\, P(R_4 \mid R_1 \cap R_2 \cap R_3)\, P(R_3 \mid R_1 \cap R_2)\, P(R_2 \mid R_1)\, P(R_1)$

$= 10/21 \;\; 12/22 \;\; 13/23 \;\; 14/24 \;\; 15/25$

$= 13/253$

$= 0.051383$

En éste ejemplo se calculó la probabilidad de un evento donde se obtienen cuatro bolas rojas en las primeras cuatro extracciones y una bola azul en la quinta extracción, sin importar cuáles sean las siguientes bolas que se extraigan para completar el experimento que consta de 10 extracciones al azar. Luego, se enumeran las extracciones y se aplica la regla de probabilidad condicional para calcular la probabilidad de que se den las cinco primeras extracciones en el orden mencionado.

La probabilidad condicional también puede utilizarse para calcular la probabilidad de eventos independientes. Cuando se habla de eventos independientes en probabilidad se quiere decir que: dos eventos *A* y *B* son *estocásticamente*[23] independientes si la probabilidad de que ocurra el evento *B* es independiente de la probabilidad de ocurrencia del evento *A*[24].

Regla de la probabilidad total

La regla de la probabilidad total es un método para calcular probabilidades utilizando la regla del producto y la regla de la suma.

---

[23] El calificativo 'estocástico' que se da a la independencia de ocurrencia de un evento *A* es necesario, pues nos referimos a que es independiente de la probabilidad de ocurrencia del otro evento *B*, y no únicamente a la ocurrencia del evento.

[24] Por Ejemplo. Si en un experimento aleatorio se tiene una urna que contiene bolas rojas y bolas blancas enumeradas, ¿cuál es la probabilidad de extraer una bola blanca en una segunda extracción? Supongamos que en la primera extracción se toma una bola roja y el experimento es con reemplazo. Si simbolizamos los eventos, tenemos:

A: en la primera extracción se obtiene una bola roja.
B: en la segunda extracción se obtiene una bola blanca.

Podemos notar que estos son eventos independientes, pues el evento *B* no es afectado si sucede *A* o no sucede:

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{\cancel{P(A)}\cdot P(B)}{\cancel{P(A)}}=P(B)$$

Ahora bien, si tenemos un experimento aleatorio[24] con eventos independientes, se calcula su probabilidad utilizando la siguiente regla que es la regla de la probabilidad condicional:

$$P(B|A)=\frac{P(A\cap B)}{P(A)}$$

En estos casos donde la ocurrencia del primer evento *A* no altera al segundo evento *B* se puede decir que la probabilidad del segundo evento es independiente de que suceda o no el primer evento, es decir, que la probabilidad del segundo evento dado el primero es igual a la probabilidad del segundo evento:

$$P(B|A)=P(B)$$

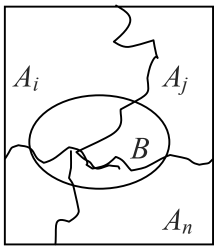
De esta forma, la probabilidad de $B > 0$ es independientemente de *A*, y se dice que dos eventos son independientes estocásticamente si la probabilidad de *B* es independiente de la probabilidad de *A*. Por lo tanto, la probabilidad de la intersección de *A* y *B* es igual al producto de estos dos eventos.

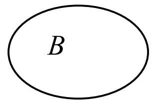
$$P(A\cap B)=P(A)\,P(B)$$

Podemos notar que la independencia estocástica afecta a la regla de la probabilidad condicional en cómo calcular los eventos de un experimento aleatorio, de esta forma, la relación que establece la independencia estocástica se impone como una condición, a cualquier modelo matemático que se quiera aplicar a un experimento aleatorio, para calcular las probabilidades de sus eventos.

Ahora bien, con la regla de la suma y con eventos mutuamente excluyentes que componen particiones del espacio muestral de un experimento aleatorio (eventos A_n); podemos calcular la probabilidad de un cierto evento B y para hacerlo necesitamos aplicar la regla de la probabilidad total:

fig.1

$\Omega=$ 

evento $B=$ 

$A_i \neq A_j$

Como notamos en la figura1, el espacio muestral Ω tiene particiones en eventos cuya unión es el total y de ello se infiere que la suma de la probabilidad de todos los eventos debe ser 1, por eso se le nombra probabilidad total. Esta es la regla de la probabilidad total:

$$\boxed{P(B)= P(B|A_1)\, P(A_1)+ \ldots + P(B|A_n)\, P(A_n)}$$

Esto queda demostrado como lo expone M.A. García:

> Los eventos $B \cap A_1,\ldots, B \cap A_n$ son mutuamente excluyentes y $B= (B \cap A_1) \cup\ldots\cup (B \cap A_n)$, así que, usando la regla de la suma primero y después la del producto, obtenemos:
>
> $P(B)= P[(B \cap A_1) \cup\ldots\cup (B \cap A_n)]$
> $=P(B \cap A_1) + \ldots +P(B \cap A_n)$
> $= P(B |A_1)\, P(A_1)+ \ldots +P(B |A_n)P(A_n)$ [García 2005, p.112]

En efecto, se suman los eventos mutuamente excluyentes y se aplica la regla del producto a cada uno de ellos[25].

[25] Por ejemplo: Se tienen dos urnas; la primera contiene 3 bolas rojas y 5 bolas negras; la segunda contiene 4 bolas rojas y 6 bolas negras. El experimento consiste en extraer al azar una bola de la primera urna y meterla en la segunda urna; la segunda parte del experimento consiste en extraer una bola de la segunda urna; con estos datos, se quiere calcular la probabilidad de obtener una bola roja en la segunda extracción.

Características que distinguen a la regla de Bayes.

Hasta ahora se han expuesto varias reglas y términos que son significativos para definir la regla de Bayes, que es atribuida a Thomas Bayes por ser el primero en utilizarla de forma explícita en su texto publicado como *de la probabilidad inversa* en 1753. La regla de Bayes actualmente es conocida como:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Lo que expresa la regla de Bayes es que al invertir el orden cuando se calcula la probabilidad condicional de un evento A dado un evento B, por la probabilidad de un evento B dado A, es más fácil calcular la probabilidad. En el cálculo de probabilidades puede suceder que se quiera calcular la probabilidad de un evento que ya sucedió, es decir, calcular la probabilidad de eventos pasados; ésto es un poco más elaborado de calcular si no se aplica la regla de Bayes; por ejemplo, si en un experimento aleatorio donde se aplica la probabilidad condicional se quiere calcular la probabilidad de un evento e_1 que antecede a otro evento e_2 que ya sucedió, y e_2 es consecuencia de e_1; este orden de causa y efecto es la desventaja de la probabilidad condicional, pues sólo permite calcular la probabilidad de

---

Si especificamos los eventos como:
A_1: se obtiene en la primera extracción una bola roja.
A_2: se obtiene en la primera extracción una bola negra.
R_2: se obtiene bola roja en la segunda extracción.
La probabilidad se calcula de la siguiente manera:

$$\begin{aligned} P(R_2) &= P\,[(R_2 \cap A_2) \cup (R_2 \cap A_1)] \\ &= P\,(R_2 \cap A_2) + P(R_2 \cap A_1) \\ &= P\,(R_2 \mid A_1)\,P(A_1) + P(R_2 \mid A_2)\,P(A_2) \end{aligned}$$

$P(R_2) = (5/11 \cdot 3/8) + (4/11 \cdot 5/8) = 35/88$

En la regla de la probabilidad total, se combinan las reglas básicas que se han visto, para abarcar más eventos que tienen diferentes características: dependientes, independientes, compuestos, simples, etc.

eventos cuando se sabe la probabilidad del evento anterior, si no se sabe, entonces es más complicado calcularla.

Un ejemplo de la aplicación de la regla de Bayes podría aclarar el problema y lo que implica:

> Imagine dos urnas, cada una contiene bolas rojas y verdes. La urna *A* tiene 80% de las bolas rojas, 20% verdes, y la urna *B* tiene el 60% verdes, 40% rojas. Escoges una urna al azar, y puedes extraer bolas de la urna, de este modo puedes preguntar qué urna es. Después de cada extracción, la bola elegida es regresada. Por lo tanto, para cualquier extracción, la probabilidad de obtener una bola roja de la urna *A* es 0.8, y para la urna *B* es 0.4. [Hacking 2001, p.71] [26]

Si definimos los datos con probabilidad condicional, obtenemos:

$P(R|B) = 0.4 \qquad P(A) = P(B) = 0.5$

$P(R|A) = 0.8$

Donde R= se extrae una bola roja.

A= se saca una bola de la urna A

B= se saca una bola de la urna B

Si extraigo una bola roja, ¿cuál es la probabilidad de que haya sido de la urna A? Es decir, ¿cuál es la probabilidad de A dado R?

$$P(A|R) = \frac{P(A)\,P(R|A)}{P(A)P(R|A) + P(B)P(R|B)}$$

$$= \frac{(0.5 \cdot 0.8)}{(0.5 \cdot 0.8) + (0.5 \cdot 0.4)} = 2/3$$

---

[26] Este es el texto original que traduje: “Imagine two urns, each containing red and green balls. Urn *A* has 80% red balls, 20% green and Urn *B* has 60% green, 40% red. You pick an urn at random, and then can draw balls from the urn in order to guess which urn it is. After each draw, the ball drawn is replaced. Hence for any draw, the probability of getting red from urn *A* is 0.8, and from urn *B* it is 0.4. You draw a red ball. What is *P*(*A*/*R*)?” [Hacking 2001, p.71]

Con la regla de Bayes es más fácil calcular $P(A|R)$ invirtiendo la probabilidad condicional , esto es, $P(R|A)$[27].

Teorema de Bayes: probabilidad *a priori* y probabilidad *a posteriori*

En la probabilidad bayesiana, los conceptos *a priori* y *a posteriori* son utilizados bajo la siguiente interpretación. Por un lado, se le nombra *a priori* a una información o conocimiento que antecede a una prueba de probabilidad que se le aplica a cierto evento, este conocimiento se puede calcular con probabilidad, por eso también se le nombra *probabilidad a priori* de un evento, y concierne de forma relevante al evento; por otro lado, se nombra *a posteriori* al conocimiento o información que resulta después de aplicar una prueba de probabilidad a un evento y se haya tomado en cuenta la probabilidad *a priori* de

[27] Otro ejemplo de probabilidad donde se utiliza la regla de Bayes es:
En un experimento aleatorio se tienen dos urnas que contienen bolas rojas y bolas negras distribuidas de la siguiente manera:
En la urna 1 se tienen 3 bolas rojas y 5 negras.
En la urna 2 se tienen 4 bolas rojas enumeradas y 6 negras enumeradas.
El experimento a realizar consiste de extraer una bola al azar de la urna 1 y meterla en la urna 2 y se extraerá una bola al azar de esta última urna.

Si simbolizamos los eventos como:

R_2: se obtiene bola roja en la segunda extracción.
R_1: se obtiene bola roja en la primera extracción.
N_1: se obtiene bola negra en la primera extracción.
Se quiere calcular la probabilidad de haber obtenido una bola roja en la primera extracción (un evento pasado), sabiendo que en la segunda extracción se obtuvo una bola roja.

$$P(R_1| R_2) = \frac{P(R_2| R_1)\, P(R_1)}{P(R_2| R_1)\, P(R_1) + P(R_2 | N_1)\, P(N_1)}.$$

$$\begin{aligned} P(R_2) &= P(R_1 \cap R_2) + P(N_1 \cap R_2) \\ &= P(R_2| R_1)\, P(R_1) + P(R_2 | N_1)\, P(N_1) \end{aligned}$$

$$P(R_1| R_2) = \frac{[5/11 \cdot 3/8]}{[5/11 \cdot 3/8] + [4/11 \cdot 5/8]} = \frac{15/88}{35/88}.$$

este evento para calcular su probabilidad. Este conocimiento *a posteriori* es factible de poder ser calculado con probabilidad.

El teorema de Bayes dentro de la teoría de la probabilidad proporciona la distribución de la probabilidad condicional de un evento A dado otro evento B que es considerado como una probabilidad *a posteriori*, porque ya sucedió; se calcula la probabilidad de A en función de la probabilidad condicional y la probabilidad del evento A que es considerada como su probabilidad *a priori*.

El teorema de Bayes tiene simetría y su probabilidad no es fija

Hay casos en donde se puede notar que hay simetría en la probabilidad bayesiana; por ejemplo: se considera que la probabilidad bayesiana tiene simetría en el sentido de que se quiere calcular $P(A|B)$ y el resultado será el mismo si calculamos $P(B|A)$, esto, sólo sucede en experimentos (con remplazo) aleatorios donde no importa si se lleva a cabo primero el evento A para calcular el evento B o si el evento B se realiza para calcular el evento A.

Otra característica de la regla de Bayes es que restringe la probabilidad de un evento, es decir, sólo calcula la probabilidad del evento en donde tiene más probabilidad de que suceda. Por ejemplo:

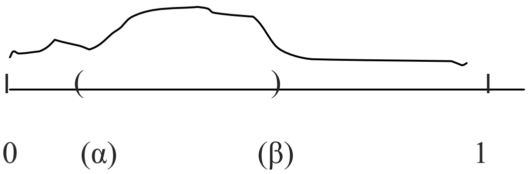
Ejemplo. Se sabe que en n observaciones el evento A ha ocurrido m veces y se quiere calcular su probabilidad, lo que hace el teorema de Bayes es calcular la probabilidad P de la probabilidad p del evento A, esta probabilidad de A es considerada como probabilidad *a priori*, pues tiene la misma probabilidad de suceder el evento A que cualquier otro evento en el espacio muestral, pero nosotros queremos saber cuál es la probabilidad de la probabilidad *a priori* de A, si sabemos la probabilidad *a posteriori* de A:

$$P\,(p(A))$$

La probabilidad de la probabilidad de A, se restringe a los experimentos realizados en donde sucedió más el evento A. En el siguiente apartado: ‘el problema de Bayes y su solución’ explicamos que Bayes propone calcular la probabilidad de la probabilidad de un evento A. Su propuesta sólo es que en vez de calcular la probabilidad de un evento A y que nos va a dar como resultado un numero x tal que $0 \leq x \leq 1$. El propone eliminar la probabilidad de que x sea igual a 0 o igual a 1, es decir, el propone que $0 < x < 1$, de esta forma restringe la probabilidad de que una probabilidad sea la probabilidad *a priori* de A. Esto presupone que hasta ahora un evento A ha sucedido varias veces en varios experimentos y podemos calcular su probabilidad. Ahora bien, la probabilidad que obtengamos es usada en la regla de Bayes para calcular la probabilidad de que suceda el evento A en nuevos experimentos que se le realizan, es decir, toma en cuenta la probabilidad anterior que tiene el evento A y esta probabilidad anterior la calcula solo fijándose en que puede ser mayor a 0 o menor que uno, es decir, $0 < P(A) < 1$

Para calcular la probabilidad de A, Bayes lo hace desde el intervalo en donde A ha sucedido más veces y, por ello, es más exacta su probabilidad; calcular la probabilidad de la probabilidad de A es calcular la probabilidad del intervalo $[\alpha, \beta]$, señalado en la figura 2; de esta forma obtenemos la probabilidad *a priori*; sabiendo la probabilidad *a priori* de A se usa como información para calcular la probabilidad *a posteriori* de A.

fig.2 El punto donde es más probable A es entre



α y β

De esta manera se calcula la probabilidad de ocurrencia de un evento, como bien explica M.A. García:

> Este resultado se debe a Bayes, quien se planteaba el problema de determinar la probabilidad de ocurrencia de un evento a partir del conocimiento del resultado de *n* repeticiones del experimento con respecto al cual está definido dicho evento. Si *p* es la probabilidad buscada, Bayes encuentra que, dado que en las *n* repeticiones del experimento el evento ocurre *k* veces, la probabilidad de que *p* esté comprendida entre dos números α y β [esta dada por la regla de Bayes]. [García 2005, p.127]

No hay que olvidar que a este tipo de casos la probabilidad de *A* no es fija, es decir, cambia al suceder más eventos; el lugar restringido por α y β cambia de posición en la línea. Si suceden más eventos, la probabilidad de *A* puede ser mayor o menor en el espacio restringido; si es menos, el espacio restringido por α y β se moverá a donde suceda más *A*.

Así pues, la probabilidad que no es fija es una de las ventajas del teorema de Bayes, pues restringe el espacio y puede cambiar la probabilidad con más información[28].

La regla de Bayes y el problema de la probabilidad inversa

La probabilidad inversa introducida por Bayes, quien retomó los estudios de Bernoulli en 1763 fue una nueva forma de calcular probabilidades de eventos distintos a los que por lo general se calculaban. Antes de la probabilidad inversa se trabajaban las probabilidades que

[28] En el fondo, en la idea de cambiar nuestra probabilidad con más información se encuentra toda una interpretación de la teoría de la probabilidad, podemos notar que en su propuesta de una probabilidad no fija, se encuentra su postura subjetiva sobre la probabilidad, es decir, cuando se refiere a una probabilidad no fija se refiere a la flexibilidad y admisibilidad de poder tratar con varias probabilidades sin restringirse a una. Como se mencionó en líneas anteriores, hay dos sentidos predominantes de entender la teoría de la probabilidad: en el sentido objetivo y en el sentido subjetivo; Thomas Bayes es considerado un representante del sentido subjetivo de la teoría de la probabilidad, pues su regla de Bayes y su postura sobre la probabilidad estuvo inclinada por la postura subjetiva de la probabilidad.

tenían resultados específicos en situaciones bien definidas, por ejemplo: se conoce cuántos lados tiene un dado, qué números tiene y su composición; en el ejemplo de las urnas, se sabe los colores de las bolas que se encuentran en las urnas, se sabe qué es una urna y cuántas urnas se usan para un experimento aleatorio, es decir, el conocimiento y la probabilidad estaban conectados. Dice Howie, refiriéndose a la investigación de la probabilidad inversa de Jacob Bernoulli:

> Jacob Bernoulli, comenzó a investigar la dirección opuesta de generar probabilidades desde los resultados observados. Conociendo la composición de las bolas en la urna, el análisis combinatorio da la probabilidad de elegir una muestra dada. Pero ¿Cómo puede uno moverse desde la evidencia de un número de elecciones a la composición de las bolas en la urna?
>
> La cuestión de la probabilidad inversa fue vista como una relación entre causa y efecto. La proporción del color de las bolas en una muestra fue causa de la razón en la urna […] Escogiendo una muestra de las bolas de la urna, uno puede calcular esta composición probable y de esta manera la probabilidad de que la siguiente bola tomada de la urna debe ser de cierto color. [Howie 2002, p.23] [29]

De esta manera, la probabilidad inversa puede ser utilizada para inferir probables causas desconocidas desde efectos conocidos. Lo que se quiere obtener con esta forma de pensar es una inversión de los métodos que usualmente se usaban alrededor del año 1750. Así pues, comenzó la búsqueda de una solución a este problema de la probabilidad inversa; estuvieron trabajando al respecto Jacob Bernoulli, de Moivre y por supuesto Thomas Bayes que, como hemos mencionado, dio una solución satisfactoria, mostró una regla y solucionó el problema.

---

[29] Este es el texto original de mi traducción: Jacob Bernoulli, they started to investigate the opposite direction, of generating probabilities from observed outcomes. Knowing the composition of balls in the urn, combinatorial analysis gives the probability of drawing a given sample. But how can one move from the evidence of a number of draws to the composition of balls in the urn?

This question of 'inverse probability'was glossed as a relation between cause and effect. The proportion of colored balls in a sample was caused by the ration in the urn, [...] Having picked a sample of balls from the urn, one could calculate its likely composition and thus the probability that the next ball drawn would of a given colour. [Howie 2002, p.23]

Según Dale, el problema que Bayes plantea y del cual parte para buscar solución es el siguiente:

> *Dado* un número de veces en el cual un evento desconocido ha sucedido y fallado: *se requiere* la ocurrencia de la probabilidad de que suceda en una simple prueba supuesta en alguna parte entre cualquiera de las dos probabilidades que pueden ser señaladas. [Dale 1991, p. 20] [30]

Es decir, dado que se ha observado varias veces que un evento desconocido a veces sucede y a veces no, conforme a estas observaciones, se quiere saber la probabilidad de que suceda el evento en una prueba que se realice en alguna parte de las probabilidades que podemos calcular, las cuales son, por un lado, la probabilidad de las veces en que ha fallado el evento y, por otro lado, la probabilidad de las veces en que ha sucedido el evento. De hecho, sabiendo la probabilidad de cualquiera de las dos podemos saber fácilmente la probabilidad del otro cálculo de probabilidad.

El problema de Bayes y su solución

Como hemos mencionado, Bayes plantea un problema que expone en uno de sus textos póstumos, este problema trata sobre la probabilidad (P) de la probabilidad (p) de un evento (A) que ya ha sucedido. Lo que se busca es: $P(p(A))$. La regla de Bayes estima un intervalo en donde puede ocurrir más el evento del que se quiere calcular su probabilidad. De forma gráfica esto lo podemos ver en la figura 2, que vimos anteriormente; en esta figura, la probabilidad de la probabilidad del evento está restringida en un intervalo y a consecuencia

[30] Este es el texto original de mi traducción: *Given* the number of times in which an unknow event has happened and failed: *Required* the chance that the probability of its happening in a single a trial lies somewhere between any two degrees of probability that can be named.[Dale 1991, p.20]

de ello se tiene un cálculo más específico de la probabilidad de la probabilidad *a priori* del evento, lo que se busca es calcular la probabilidad *a priori* de un evento en específico[31].

Hasta ahora se ha desarrollado el teorema de Bayes en el sentido formal de la probabilidad, el sentido que se le ha dado al construir un modelo matemático para calcular la probabilidad de los eventos que expresan propiedades objetivas; por ejemplo, cuando dice que se tiene un dado (no trucado) y se quiere calcular la probabilidad de obtener el número 5, se sabe que se tiene equiprobabilidad, una probabilidad fija, pues el dado sólo puede dar 6 posibles resultados. Se comienza con evidencia que no cambia, que sabemos y damos por hecho, el dado siempre va a tener 6 lados. Pero como se ha dicho, este no es el único enfoque de calcular probabilidades; también está el enfoque subjetivo en el que Bayes trabajó. En la regla de Bayes podemos distinguir las características de una interpretación subjetiva de la probabilidad, que la probabilidad objetiva no aceptaría; por ejemplo, la flexibilidad de poder cambiar de probabilidad con más información. Esta es una característica de la probabilidad subjetiva, la probabilidad que se considere ahora sobre un evento no será la misma después de que hayan sucedido varios experimentos sobre el mismo evento, es decir, el conocimiento que obtengamos de los experimentos afectará a la probabilidad del evento que ya habíamos calculado.

La probabilidad inversa, ahora también conocida como la regla de Bayes, cambia la probabilidad condicional buscada, es decir, si se busca la probabilidad de un evento *B* dado un evento *A*, se invierte la búsqueda de probabilidad: se busca la probabilidad de un evento *A* dado un evento *B*. Lo que se hizo fue invertir la búsqueda de probabilidad, este cambio o

[31] Por seguimiento al objetivo de este texto, no nos detendremos en la solución formal del problema de la probabilidad inversa, pues se requiere varias nociones técnicas en probabilidad y fórmulas de probabilidad; pero se puede ver la solución formal del problema en [Dale 1991, pp. 20-29].

inversión es muy útil, pues si se tiene la probabilidad de uno de los eventos se puede calcular la del otro evento ya sea el consecuente o el antecedente de la probabilidad condicional, esto es, se puede calcular la probabilidad de las causas desconocidas desde efectos conocidos.

En el campo científico las dos posturas probabilísticas que más se distinguen son la interpretación objetiva y la interpretación subjetiva de la probabilidad; en estas interpretaciones la mayoría de las teorías de la probabilidad sólo son aplicables a eventos, si cumplen ciertas características que tal o cual interpretación exige para poder calcular la probabilidad de tal evento. Por ejemplo, la teoría frecuencialista de la probabilidad es una interpretación objetiva de la probabilidad y sólo se aplica a eventos que cumplen las nociones objetivas de la probabilidad. En cambio, la teoría bayesiana de la probabilidad tiene su origen en la interpretación subjetiva de la probabilidad y, sin embargo, la regla de Bayes no necesita que un evento cumpla con las nociones subjetivas de la probabilidad para poder calcular su probabilidad con el teorema de Bayes; la regla de Bayes puede ser aplicada tanto en una interpretación subjetiva como en una objetiva de la probabilidad; por ejemplo, en el texto de Galavotti cita al científico geofísico Harold Jeffreys (1891-1989) quien asumió el método Bayesiano como una herramienta importante en el método científico para seleccionar hipótesis, después de haber usado la teoría frecuencialista:

> Jeffreys asumió el método Bayesiano en su trabajo como científico. Trabajando en el campo de la geofísica, […] en su experiencia como practicante de la ciencia constantemente estaba enfrentándose con problemas de probabilidad inversa, tenía que explicar datos experimentales para significados de diferentes hipótesis, o para evaluar hipótesis generales a la luz de datos cambiantes. […] la ciencia representa una rama del tema de la probabilidad. De hecho, de acuerdo a Jeffreys la probabilidad debe ser considerada como aquella que nos permite tener un mayor alcance del amplio campo de la ciencia, porque el proceso de adquirir conocimiento se relaciona con la probabilidad en los primeros pasos. […]Con las bases del Bayesianismo encontramos el supuesto de que toda ley cuantitativa forma un sistema numerable, y sus probabilidades forman una

> serie convergente. […]Tiene en cuenta la asignación significativa de probabilidades *a priori* a hipótesis generales. En resumen, Jeffreys y Wrinch formularon un 'postulado de simplicidad' que dice las leyes simples son asignadas a una gran probabilidad *a priori*. [Galavotti 2005, pp. 179-181] [32]

En el tema de la aplicación del teorema bayesiano a la teoría de la corroboración de hipótesis científicas se verá este testimonio con más detalle.

1.3 La axiomatización de la probabilidad

Tanto la vertiente cualitativa que también, como mencionamos, es conocida como la probabilidad subjetiva, y la vertiente cuantitativa que también, como mencionamos, es conocida como la probabilidad objetiva son interpretaciones de la probabilidad; sin embargo no cualquier interpretación de la probabilidad es legítima. A pesar de que no hay una definición única de la probabilidad, sí hay axiomas que deben satisfacer las interpretaciones de la probabilidad para que sean consideradas como rigurosas. Estos axiomas fueron propuestos por Kolmogorov (1950) en su documento *Fundamentos de la teoría de la probabilidad* y son actualmente conocidos como la terna fundamental para modelar un fenómeno. Según Miguel Ángel García:

> El modelo matemático de un fenómeno probabilístico está dado por una terna $(\Omega, \gimel, P)$, en donde Ω es un conjunto, $\gimel$ una σ – álgebra de subconjuntos de Ω y P una medida de probabilidad definida sobre $\gimel$. [García 2005, p.384]

---

[32]Este es el texto original de mi traducción: Jeffreys was led to embrace Bayesian method by his own work as a scientist. Working in a field like geophysics [...] Moreover, in his experience as a practising scientist he was typically faced with problems of probability inverse, having to explain experimental data by means of different hypotheses, or to evaluate general hypotheses in the light of changing data. [...] In fact, according to Jeffreys probability should be seen as having a broader scope than science, because the process of acquiring knowledge is permeated with probability from its first steps. [...] At the basis of the Bayesianism of Jeffreys and Wrinch we find the assumption that all quantitative laws from an enumerable set, and their probabilities from a convergent series. [...] allows for the assignment of significant prior probabilities to general hypotheses. In addition, Jeffreys and Wrinch formulate a 'simplicity postulate', according to which simpler laws are assigned a greater prior probability.[Galavotti 2005, p.179-181]

Así pues en la terna presentada por Kolmogorov, (Ω, *P*, ℷ) podemos ver que *P* es una función de probabilidad, Ω un espacio que es igual a 1 en el sentido de que todos los posibles resultados de probabilidad se encuentran en Ω, y los posibles resultados de calcular la probabilidad es un número entre cero y 1. ℷ es un conjunto de subconjuntos de Ω. Tanto el método frecuencialista[33] de la probabilidad como el método bayesiano de la probabilidad satisfacen los axiomas que en el campo de la probabilidad se consideran esenciales en toda teoría o método de probabilidad que aspire a ser considerada como rigurosa y formal.

Conclusión

En este primer capítulo llegamos a la conclusión de que las definiciones que se han dado de la probabilidad a lo largo de la historia son interpretaciones que tratan de ser más precisas en abordar los problemas que se tienen desde los juegos de azar hasta los problemas de toma de decisiones en la vida diaria.

En el caso de la probabilidad clásica, Laplace supone que todos los eventos son equiposibles, pero surge otra teoría de la probabilidad que afirma que los eventos no son equiposibles (probabilidad clásica), por ejemplo, el juego de los volados, los eventos no pueden ser equiposibles, pues la evidencia de eventos pasados, es decir, si antes ya se ha lanzado la moneda varias veces se espera que la frecuencia ayude a calcular la probabilidad; en cambio, un evento que es equiposible es aquel en el que no se puede hacer experimentos que le antecedan al evento que queremos calcular su probabilidad, por ejemplo, queremos saber la probabilidad de que al dispararle a un gato en específico éste no se muera en los primeros disparos, en este caso la probabilidad de que el gato no se

[33] Para ver la prueba formal, puede remitirse al texto de [Gillies 2000, p.59-112]

muera en los primeros disparos es equiposible en el sentido de que puede y no puede morir el gato y no podemos decir que morirá, pues con una probabilidad que se basa en la experiencia en la frecuencia no la tenemos.

Por otra parte la definición de probabilidad frecuencialista considera que todo evento tiene una frecuencia. Los intentos por mejorar la teoría de las probabilidades, que tratan de describir y estudiar fenómenos de la naturaleza, en algunos casos se enfocan más en los problemas de cálculo; otros en el método que aplican. En cualquiera de los casos, los principios y las definiciones resultantes intentan dar un método sólido que es esencial en mucho del conocimiento científico.

Hasta ahora, las interpretaciones de probabilidad que hemos visto sólo se basan en experimentos aleatorios, pero hay experimentos que no se pueden repetir bajo las mismas condiciones; por ejemplo:

> Tomemos un gato para experimentar, […] el cual está completamente sano y no tiene ninguna herida. Consideremos entonces que el experimento consiste en elegir una persona de manera arbitraria, darle una pistola cargada y pedirle que dispare sobre el gato. Como resultado del experimento, observamos si, después de realizarlo, el gato está vivo o muerto. Si quisiéramos considerar una serie de repeticiones de este experimento, en un momento dado, posiblemente después de su primera realización, este experimento ya no será repetible, pues el gato ya estará herido o muerto. [García 2005, p.33]

Así pues, si el gato está herido o muerto, ya no podemos repetir el experimento, pues, las condiciones originales en que se ha realizado el experimento ya no se pueden reproducir. En este ejemplo y otros parecidos son por lo general más complicados, pues limita la posibilidad de poder repetir los experimentos.

También se presentan dificultades cuando los factores subjetivos no son considerados como parte del método probabilístico, pues afecta de manera considerable el resultado; vimos en el tema 'el uso de la probabilidad epistemológica y la evidencia' un ejemplo de Ian Hacking que citamos sobre esta dificultad.

En la primera parte del ejemplo de Hacking, sobre los contratos debidamente notariados, tenemos una probabilidad cuantitativa, pero los datos siguientes del ejemplo introducen un elemento cualitativo y esto afecta la probabilidad cuantitativa de una forma cualitativa reduciéndola hasta eliminarla. Este es el problema de la evidencia mixta donde una parte apoya la hipótesis y otra parte la contradice con datos cualitativos; el ejemplo nos muestra, en otras palabras, que hay casos en donde la probabilidad cualitativa debilita la hipótesis de una probabilidad cuantitativa.

Hasta ahora sólo hemos visto interpretaciones probabilísticas cuantitativas, pero en el ejemplo anterior nos muestra factores cualitativos que no se tenían contemplados y que son de relevancia para el resultado que dé la teoría de la probabilidad. En casos como el anterior, parece que es necesario un punto de vista subjetivo donde se consideren eventos que no se repiten, es decir, la subjetividad y creencias individuales sobre el suceso a ocurrir. Tales factores son considerados en las interpretaciones de la teoría de la probabilidad como cualitativos; de esta forma la probabilidad cualitativa incorpora un concepto subjetivo que expresa la confianza individual que se tiene sobre la ocurrencia de tal evento. Una definición de probabilidad que retoma estos factores subjetivos, es la definición de probabilidad bayesiana.

II. Interpretaciones filosóficas de la probabilidad

Introducción

En este capítulo se expondrá la interpretación que Russell da a la teoría de la probabilidad, esto se hará por razones históricas y porque es pertinente explicar por qué la manera en que se interpreta la probabilidad puede incidir en su aplicación, en particular puede causar que se obtengan distintos resultados y se justifiquen diferentes inferencias tras la aplicación de métodos probabilísticos. Destacaré dos interpretaciones de la probabilidad, a saber, la interpretación subjetiva y la objetiva, evaluando algunos problemas y ventajas de cada una. También expondré cómo es que la interpretación de la probabilidad que asumen los científicos afecta en las inferencias que realizan así como el proceso de corroboración o refutación de hipótesis.

2.1 Probabilidad y grados de credibilidad

A continuación expondremos las propuestas que surgieron cuando Russell se dio cuenta que la probabilidad significaba más que reducir los hechos de la vida real o la ciencia a números. La aplicación de la probabilidad es un proceso más complicado de lo que parecía; Russell, entre otras personalidades que influyeron en la teorización filosófica de la probabilidad, relacionó la probabilidad con las creencias y el conocimiento que estaba de alguna forma implícito en ella. La aportación de Russell a la teoría filosófica de la probabilidad consiste en la conceptualización de grados de credibilidad que supone que afectan en la justificación de nuestro conocimiento. Keynes, como Russell, apoya la teoría de los grados de credibilidad, aunque con algunas diferencias.

En este subcapítulo, primero, daremos una exposición de las interpretaciones de la probabilidad, después las diferentes clasificaciones que se hacen sobre estas interpretaciones; luego se expondrá la relación que hay entre nuestras creencias y conocimiento, de inmediato se mostrará la definición de Russell de lo que es un *dato*. Luego, explicaremos qué son los grados de credibilidad para Russell y como los usa para sostener que la teoría matemática tiene como último motivo de orientación para saber qué concepto de probabilidad (que satisface los axiomas) debe ser considerado como aceptable. Posteriormente se expondrán las críticas que Russell sostiene sobre los orígenes de nuestras creencias (la percepción, la memoria y la inferencia). Por último se profundizará sobre la distinción entre teoría subjetiva y objetiva de la probabilidad.

Interpretaciones de la probabilidad

Russell en su texto *'Human Kowledge'* publicado en (1930/1992), propone hacer un análisis de los diferentes conceptos de probabilidad. El primero que trata es la teoría matemática de la probabilidad; se pregunta por la explicación que dan los matemáticos que coinciden con esta teoría, para argumentar a favor de ella. Si bien por un lado los matemáticos coinciden en que puede ser expresada la teoría en términos matemáticos; por otro lado, no hay un acuerdo en cómo debe ser interpretada. En tales casos, lo que hacen los matemáticos es simplemente decir que cualquier concepto que se use para interpretar la teoría matemática que pueda satisfacer los axiomas de la probabilidad, se puede considerar como una interpretación rigurosa y formal; por ejemplo, Russell nos muestra un concepto de probabilidad que satisface los axiomas de la teoría de la probabilidad: "Dada una clase finita B la cual tiene n miembros, y dado que m de estos pertenecen a alguna otra clase A, entonces decimos que, si un miembro B es elegido al azar, la probabilidad de que éste pertenezca a la clase *A* es *m/n*."[34] Este concepto satisface los axiomas, pero como éste puede haber otros (por ejemplo, los conceptos pertenecientes a la teoría de Bayes que también los satisfacen), por ello, para Russell los motivos que se tienen para elegir conceptos está fuera de las matemáticas, pues cualquier concepto que pueda satisfacer los axiomas puede ser tomado como aceptable para una definición de la teoría matemática de la probabilidad. Como dice Russell: "De hecho, puede ser conveniente adoptar una interpretación en un contexto, y otra en otro, la conveniencia es el único motivo de orientación."[35]

De esta forma, la interpretación que se elija para la teoría matemática de la probabilidad puede depender, en última instancia, del uso que se le quiera dar, o mejor dicho de lo que nos conviene usar. Un problema que se origina al no tener una interpretación única de la teoría matemática de la probabilidad es que dos interpretaciones distintas pueden satisfacer los axiomas pero arrojar diferentes probabilidades, ambas serán legítimas desde el punto de vista matemático; por ejemplo:

> Sabemos que el evento en cuestión es uno de los eventos de la clase *B* y podemos saber qué proporción de esta clase pertenece a una clase *A* de la cual estamos interesados. Pero la proporción variará de acuerdo a nuestra selección de la clase *B*; de esta manera, podemos obtener distintas probabilidades, sin embargo todas son igualmente válidas desde el punto de vista matemático. [Russell 1992, p.358] [36]

En otras palabras, dos métodos igualmente válidos con respecto a la teoría axiomática de la probabilidad pueden ser empleados para calcular la probabilidad de cierto

---

[34] Este es el texto original de mi traducción: Given a finite class *B* which has *n* members, and given that *m* of these belong to some other class *A*, then we say that, if a member of *B* is chosen at random, the chance that it will belong to the class *A* is *m/n*. [Russell 1992, p.356]

[35] Este es el texto original de mi traducción: In fact, it might be desirable to adopt one interpretation in one context, and another in another, for convenience is the only guiding motive. [Russell 1992, p.357]

[36] Este es el texto original de mi traducción: We know that the event in question is one of a class *B* of events, and we may know what proportion of this class belongs to some class *A* in which we are interested. But the proportion will vary according to our choice of the class *B*; we shall thus obtain different probabilities, all equally valid from the mathematical standpoint. [Russell 1992, p.358]

evento y entregar resultados distintos, debido a que el proceso de selección de datos, que afecta el resultado, no es relevante desde el punto de vista de los axiomas de la probabilidad. Sin embargo, es evidente que para el quehacer científico sí es importante el proceso de selección de datos, pues de este proceso dependerán los resultados que afectarán en la corroboración o refutación de hipótesis.

Ahora bien, la 'probabilidad' no sólo puede ser interpretada desde un punto de vista matemático, sino que en los métodos científicos frecuentemente es usada para corroborar hipótesis de una forma que parece muy poco obvia desde el punto de vista matemático. Russell destaca dos interpretaciones no matemáticas de la 'probabilidad'. La primera había sido enfatizada por el obispo Butler (1736), quien afirmada que la probabilidad es la guía de la vida[37]. La segunda es enfatizada por Reichenbach, él decía que todo nuestro conocimiento es sólo probable[38]. Para Russell la primera frase, propuesta por Butler, es aceptada por cualquier interpretación de la probabilidad, pues; por ejemplo, nuestro conocimiento es limitado cuando estamos en la situación de elegir una hipótesis de varias, Russell dice que por lo general elegimos la más probable. Por otro lado, cuando decimos, como Reichenbach, que todo nuestro conocimiento es sólo probable, estamos diciendo que nuestro conocimiento es limitado; lo que debemos enfatizar es que la probabilidad en la práctica tiene una fuerte relación con la credibilidad al igual que la convicción subjetiva, ambas están relacionadas y pueden ser estudiadas empíricamente.

Russell afirma que estas dos interpretaciones no matemáticas pueden adaptarse a la teoría matemática de la probabilidad (axiomática), pero no son consideradas como

37 [Butler en Hacking 1995, p.107]
38 [Reichenbach en Hacking 1995, p.107]

relevantes en esta teoría, sino que son supuestas y aceptadas sin examen. Debemos reconocer que vinculan una interpretación de la probabilidad la cual parece, según Russell, imposible de reconciliar con la definición de la teoría de la probabilidad matemática.

De esta forma, las distintas interpretaciones de probabilidad que vinculan distintas definiciones de la 'probabilidad' nos llevan a sugerir una pregunta: ¿qué es lo que debemos considerar por probabilidad? ¿El estudio de los datos a los que se les aplica la probabilidad o sólo el método que se utiliza para calcular la probabilidad? Aunque creemos que la última objeción debe de ir dirigida a la estadística más que a la probabilidad, pues la elección de los datos es objeto de estudio de la estadística, sin embargo, la distinción entre qué preguntas le corresponde a la probabilidad y a la estadística se elimina con el método bayesiano, pues este método puede tratar con el estudio de los datos y el cálculo de la probabilidad.

En resumen, Russell expone dos diferentes conceptos de la probabilidad, cada uno de ellos se distingue por tratar con eventos que tienen distintas características, esta distinción es la que nos interesa desarrollar. El primer concepto es el dado por la teoría matemática de la probabilidad que se distingue por utilizar sólo casos que pueden ser numerables, que satisfacen los axiomas del cálculo de probabilidad y se restringe al estudio de colecciones y no de casos simples. El segundo concepto es el de 'grados de credibilidad', este concepto de probabilidad se aplica a proposiciones simples y siempre toma explicaciones de toda la evidencia relevante. Este último concepto es el que desarrolla y apunta. Hasta ahora Russell expone dos intentos de definir 'probabilidad'. Como dice Russell: "Una de las causas de las teorías fallidas ha sido el fracaso de distinguir – o mejor dicho, la determinación para confundir – esencialmente conceptos diferentes que, son usados, tienen un igual derecho de ser llamados ‘probabilidad’”[39]

Clasificaciones de las teorías de la probabilidad

La idea de probabilidad por lo general conduce a dos concepciones diferentes: probabilidad subjetiva (creencia) y probabilidad objetiva (frecuencia). Por un lado, la probabilidad puede hacer que pensemos en el grado por el cual podemos sentirnos confiados en algo que es incierto desde lo que sabemos o podemos averiguar; por otro lado, la probabilidad también puede hacer que pensemos en la frecuencia relativa en donde ocurre algún resultado en pruebas repetitivas sobre una estructuración de la oportunidad o lo aleatorio. Sin embargo, ambas teorías de la probabilidad pueden ser legítimas, aun siendo distintas entre ellas filosóficamente y en su utilidad, pues en el lenguaje ordinario usamos la palabra ‘probabilidad’ de diferentes formas, ya que no hay un único concepto de probabilidad, pero para que éstas teorías sean consideradas legítimas sí deben de satisfacer los axiomas de la teoría axiomática de la probabilidad.

En la actualidad, las teorías de la probabilidad más conocidas son la subjetiva y objetiva (creencia y frecuencia, respectivamente), también son conocidas como la probabilidad cualitativa y cuantitativa. Hemos de notar que esta gran distinción de interpretaciones de la probabilidad es tajante. Por un lado la interpretación subjetiva de la probabilidad tiene una justificación de orden epistemológico; ésta afirma que las probabilidades están relacionadas con las personas y las medidas del conocimiento humano o creencias. Por otro lado, la teoría objetiva de la probabilidad afirma que la probabilidad es

---

[39] Este es el texto original de mi traducción: One of the causes of faulty theories has been failure to distinguish –or rather, determination to confound- essentially different concepts which, so far as usage goes, have an equal right to be called “probability”. [Russell 1992, p.356]

el cálculo de frecuencias en que ocurre un evento *x,* y no está influenciado por ningún factor subjetivo.

De las dos distintas interpretaciones de la probabilidad, los matemáticos creen que la definición de la probabilidad objetiva es clara y sin ambigüedades, pero no dice nada sobre el uso particular de la probabilidad en los métodos científicos; en cambio, la teoría subjetiva de la probabilidad, dice algo al respecto del uso de la probabilidad en los métodos científicos, pero a cambio de ampliar más su definición de probabilidad pierde claridad y formalidad, más adelante veremos esto con más detalle.

Así como hay distintas interpretaciones de la probabilidad, también hay distintas formas de clasificar los tipos de teorías de la probabilidad; por ejemplo, Wesley Salmon dice que "la probabilidad debe ser dividida como objetiva (mundo material) y epistemológica (conocimiento humano)"[40]; Ian Hacking sugiere llamar a este último grupo epistemológico como todas aquellas teorías que se identifican con grados de creencia, de creencia racional, conocimiento o ignorancia. "Por un lado es estadística y le concierne a las leyes estadísticas. Por otro, es epistemológica, dedicada a estimar grados razonables de creencia"[41]. En cambio, Gillies considera que: "las interpretaciones epistemológicas de la probabilidad son apropiadas para la economía y las ciencias sociales, mientras que las ciencias naturales requieren una interpretación objetiva de la probabilidad"[42].

Grados de credibilidad

---

40 [Salmon 1984, p.57]

41 [Hacking 1995, p.25]

42 Este es el texto original de mi traducción: […]epistemological interpretations are appropiate for economics and the social sciences, whereas the natural sciences require an objective interpretation of probability. [Gillies 2000, p.3]

Ahora bien, de las interpretaciones y clasificaciones que puede haber sobre la probabilidad Russell apoya aquella que habla sobre grados de creencia, él afirma que por lo general tenemos un cierto grado de credibilidad[43] en nuestro conocimiento, este grado de credibilidad puede cambiar; por ejemplo, hoy podemos estar seguros de un conocimiento, pero mañana puede cambiar la confianza[44] que tenemos sobre tal conocimiento cuando tengamos más datos. Pero, ¿cómo es posible que cambie el grado de credibilidad de algún conocimiento que tenemos? Russell dice que cuando argumentamos para sustentar una creencia[45] queremos apoyarla con razones y nos apoyamos de evidencia que puede aumentar su grado de credibilidad o puede disminuirlo; para saber cuándo aumenta o disminuye su grado de credibilidad, hay varios factores, circunstancias que influyen y que deben ser interpretadas con un criterio que la lógica deductiva no nos puede dar, pues ella trata sólo con dos valores: una proposición o es verdadera o es falsa. Nosotros estamos hablando de grados de credibilidad, una escala entre la verdad y la falsedad. Lo que nos serviría es la probabilidad, según Russell:

> Hay una cierta conexión entre la probabilidad matemática y los grados de credibilidad. Esta es: cuando en relación para toda evidencia disponible, una proposición tiene cierto grado de probabilidad, entonces esta es la medida del grado de credibilidad. [Russell 1992, p. 339]

La probabilidad matemática nos puede ayudar para saber qué grado de credibilidad debemos dar a tal creencia, pues sólo traduce en números la predicción de un evento, la

---

43 Hemos de destacar que al referirnos al concepto de grados de credibilidad estamos haciendo referencia a una característica esencial de la teoría subjetiva de la probabilidad. También debemos de distinguirla del concepto grados de credibilidad racional que es un concepto fundamental y una de las características que hace distinguir a la teoría lógica de la probabilidad, si el lector desea saber más sobre las distintas teorías de la probabilidad puede consultar [Gillies 2000, pp.51-87].

44 Para Russell la confianza se puede interpretar como el grado de credibilidad que tenemos sobre *x* conocimiento.

45 La creencia en este texto es definida como el grado de credibilidad que nosotros tenemos sobre cierto conocimiento.

repetición del experimento puede reducir el riesgo de error hasta que llegue a ser mínimo; de este modo puede aumentar el grado de credibilidad. Todo esto está dentro del campo de la teoría matemática de la probabilidad, pero sólo nos puede ser útil cuando es aplicable, es decir, habrá casos en que no van a ser sólo cuantitativos sino que también pueden tener algunos elementos cualitativos, como el ejemplo de los notarios citado de Hacking en el primer capítulo de este texto. Russell afirma que el grado de credibilidad al que se refiere puede ser aplicable más allá de la probabilidad matemática, se puede aplicar a casos que son cualitativos.

También, Russell cree que a cualquier proposición que sostengamos se le puede aplicar grados de credibilidad, pero estas proposiciones deben de estar relacionadas con datos que afecten su grado de credibilidad, ya sea que les favorezcan o no. "Si lo que digo es lógicamente aceptable, debemos creer que el grado de credibilidad que se relaciona con una proposición es esto mismo considerado como un dato."[46]

Ahora bien, Russell indica la complejidad de saber el grado de credibilidad de un argumento, pues, un argumento relaciona las premisas de cierta forma que trata de hacer que la conclusión se siga de ellas, pero es difícil saber esto, si no se es un experto en el tema que se desarrolla en el argumento o si es un argumento demasiado grande. Por ello, Russell cree que para estar seguros que una conclusión se sigue de una premisa, debemos 'ver' cada paso del argumento y descomponer el argumento en pasos más sencillos. Por ello, Russell está de acuerdo con lo que menciona Hume al respecto:

> En tales argumentos, como Hume precisa, la conclusión de un gran argumento tiene menos certeza que la conclusión de uno corto, porque en cada paso hay un cierto riesgo

[46] [Russell 1930, p. 402]

> de error. [Un argumento que está compuesto de más pasos tiene más riesgo de error que un argumento que está compuesto de menos pasos.] [Russell 1992, p. 400][47]

Pero, si bien la probabilidad matemática nos da evidencia numérica que podemos usar para dar un cierto grado de credibilidad a tal proposición, no toma en cuenta la convicción privada.

> Lo que, sin embargo, no está dentro del ámbito de aplicación de esta teoría es la convicción privada [personal] del individuo como el matemático que toma en cada paso; pero a pesar de esta variabilidad debe ser lo más directa e inmediata como nuestra confianza en los objetos de la percepción. [Russell 1992. p. 400][48]

La convicción privada o personal puede afectar de manera decisiva la conclusión de un argumento de la clase de inferencias científicas que sólo son probables; Russell llega a decir que las inferencias de las ciencias son distintas a las de la lógica y de la matemática, porque son probables, esto es, cuando las premisas son supuestamente verdaderas y el argumento correcto, la conclusión sólo puede ser probablemente verdadera; en este caso, esto sucede porque influyen varios factores, como la convicción privada que afecta en el argumento y es representada por el grado de credibilidad. Por ejemplo, cuando un argumento científico es estudiado por un maestro y su alumno, para saber si es correcto; el alumno sabe que el maestro es un experto en el tema que se desarrolla en el argumento y también que el maestro ha afirmado que el argumento es correcto, con estos datos el grado de credibilidad de que el argumento sea correcto ha aumentado en el alumno, por consiguiente afirma que si las premisas son verdaderas, el argumento válido y el grado de probabilidad es muy alto por los datos que él sabe por experiencia personal, entonces concluye que el argumento es correcto y que la conclusión es muy probablemente verdadera. En este ejemplo, el alumno infiere datos que relaciona con creencias o reflejos

[47] Este es el texto original de mi traducción: On such grounds, as Hume points out, the conclusion of a long argument has less certainty than the conclusion of a short one, for at each step there is some risk of error.
[48] Este es el texto original de mi traducción: What, however, is not within the scope of that theory is the private conviction of the individual mathematician as he takes each step; but in spite of this variability it must be as direct and immediate as our confidence in objects of perception. [Russell 1992. p. 400]

condicionados por la costumbre, aunque acepta que hay en el fondo una estructura lógica deductiva. Los datos que se infieren de esta forma son los que afectan el grado de probabilidad en una inferencia científica.

Datos para los grados de credibilidad

Un dato, es lo que constantemente Russell refiere como la evidencia que puede aumentar o disminuir el grado de credibilidad de una proposición, por ello, se pregunta qué debemos considerar como dato y cómo afecta la proposición de la cual queremos saber su grado de credibilidad. Russell afirma que el grado de credibilidad[49] se aplica a proposiciones y los datos son expresados en proposiciones; si un dato es considerado como seguro puede incrementar el grado de credibilidad de un argumento.

La relación epistemológica entre un dato y las proposiciones inferidas puede llegar a complicarse; por ejemplo, en la recolección de datos podemos pensar que estamos recolectando datos para apoyar una hipótesis, es decir, aumentar el grado de credibilidad, pero ¿cómo justifico que estoy recolectando datos? O ¿Cómo sé qué son datos? Las respuestas que se den serán ambiguas, por consiguiente tengo razones para rechazar los datos como evidencia a favor de la hipótesis; Russell trata de clasificar los datos que influyen en nuestros grados de credibilidad.

Russell distingue datos que pueden ser inferidos desde otros datos, estos son los que concluimos desde los argumentos, las premisas de los argumentos las considera como

[49] Russell hace referencia a la distinción de grados de credibilidad y grados de credibilidad racional; el segundo concepto no se debe confundir con la teoría lógica de la probabilidad, pues Russell, cuando habla de grados de credibilidad racional sólo se refiere a lo que define como dato y define dato como una proposición la cual tiene un grado de credibilidad racional en su explicación; el grado de creencia racional sólo es dado por las creencias racionales, que, como dice Russell pueden tener tres distintos orígenes, más adelante veremos estos orígenes.

datos; también distingue datos que son heredados y datos que pueden ser dados por la memoria, estos últimos datos los considera como no concluyentes, es decir, no afectan de forma decisiva el grado de credibilidad. Por ejemplo, si no tengo buena memoria para recordar a qué hora cené con una persona específica el año pasado, cualquier conjetura que dé al respecto no será confiable, pero si tengo anotada la hora y la fecha en que cené con la persona, puedo corroborar concluyentemente que sé a qué hora cené el año pasado con tal persona. De esto se sigue que cualquiera de mis creencias pueden ser debilitadas o reforzadas por la evidencia que tengo.

En efecto, la memoria no proporciona datos confiables para apoyar una hipótesis, pero sí afecta mi grado de credibilidad sobre tal hipótesis; la memoria puede fallarnos y recordar que cené con alguien el año pasado a tal hora, puede cambiar de forma radical cuando tengo otras creencias más sólidas que apoyan o refutan la razón que me da la memoria para sustentar la hipótesis; el haber escrito a qué hora cené con esa persona puede afectar las razones que tengo para apoyar una hipótesis, de esta forma aumenta o disminuye mi grado de credibilidad. Hay creencias que pueden afectar de forma decisiva otras creencias. El hecho de que recuerde una ocurrencia puede ser evidencia, pero no una evidencia concluyente. Russell llama evidencia ocurrente a la que no afecta de forma decisiva mi creencia de que yo cené a tal hora el año pasado, esta evidencia es la que me da la memoria; en cambio, llama evidencia concluyente a la que sí afecta de forma decisiva mi creencia; ésta puede ser el encontrar anotado en una hoja la hora y fecha en que cené con aquella persona el año pasado. De esta forma, Russell relaciona las creencias que tienen distintos grados de credibilidad y que afectan de forma ocurrente y de forma decisiva el grado de credibilidad que tengo sobre tal hipótesis.

Los datos concluyentes son premisas que tienen un grado de credibilidad racional por sí misma, es decir, si relacionamos otras proposiciones con la premisa, no será afectado su grado de credibilidad; por ejemplo, remitiéndonos al ejemplo anterior, la memoria no sería una premisa, en cambio, el escrito de la hora y fecha en que cené sí sería una premisa. En esto se basa la ciencia de hechos no de ocurrencias, por consiguiente a pesar de que no son hechos cuantitativos sino que son cualitativos pueden ser relevantes para apoyar una hipótesis dada la evidencia; en este caso es evidencia con características cualitativas que sólo los métodos probabilísticos con posturas subjetivas pueden tratar, sin embargo la mayoría de los métodos de la probabilidad sólo tratan o eventos con características cuantitativas o cualitativas, pero no ambos. El método bayesiano de la probabilidad trata con ambos tipos de eventos por ello considero que es un método de la probabilidad con ventajas en su uso en los procesos científicos que corroboran o refutan hipótesis, pues se puede usar tanto en métodos de probabilidad objetivos y subjetivos. Ahora bien, en los métodos científicos la mayoría maneja información tanto cualitativa como cuantitativa, por consiguiente lo mejor sería usar un método de la probabilidad que maneje ambos tipos de información. Aunque los métodos científicos intentan ser objetivos, no es posible que lo sean por completo. Uno de los factores que influyen en su falta de objetividad es el manejo de los eventos como la evidencia que no puede evitar el científico usar. A pesar de que la evidencia involucra varias dificultades con características subjetivas (evidencia con características cualitativas) el científico busca métodos (como el método bayesiano) que les dé rigor cuando manejan creencias que son originadas de la interpretación de la evidencia. Sin embargo, puede interpretarse que el método bayesiano realmente no está siendo objetivo cuando maneja la evidencia y se puede tomar como un defecto de este método, más adelante veremos con detenimiento el problema de la interpretación de los eventos como evidencia.

Ahora bien, Russell considera datos como creencias que heredamos de las personas que ya han experimentado; éstas son las creencias que tomamos del resultado del esfuerzo de otras personas. Distingue dos tipos de credibilidad como la credibilidad intrínseca que nosotros formamos y la credibilidad que es heredada, aunque cualquier credibilidad intrínseca puede ser originada de alguna otra credibilidad que sea considerada como heredada. Después de distinguir los distintos tipos de datos, Russell afirma que el cálculo de la probabilidad puede ser interpretado desde grados de credibilidad.

Credibilidad en la probabilidad matemática

El cálculo de la probabilidad afecta el grado de credibilidad que tenemos sobre una hipótesis; por ejemplo, al tomar una carta al azar de una baraja completa de cartas, el grado de credibilidad de que la carta tomada sea de color rojo es exactamente igual al grado de credibilidad de la carta que no es roja, pues el grado de credibilidad es ½. Por ello, todas las frecuencias derivadas de la teoría de la probabilidad clásica pueden ser interpretadas como grados derivados de la credibilidad.

Lo que hace la teoría matemática de la probabilidad es explicar casos; pero al traerla a los grados de credibilidad tenemos que saber o asumir que cada caso es igualmente creíble. Russell afirma que cada caso o hecho probable que tengamos tiene el mismo grado de credibilidad, pero la probabilidad no sólo se limita a la probabilidad clásica, sino que ha ido evolucionando para tratar no sólo con números sino con grados de credibilidad, esto último se hace en la postura subjetiva de la probabilidad.

Ahora bien, Russell dice que los hechos con los que trabaja la teoría matemática de la probabilidad son restringidos, pues todos los datos usados para aplicar la teoría matemática de la probabilidad comienzan desde una colección bien definida; por ejemplo, una bolsa de bolas, un paquete de cartas. El punto es que las frecuencias siempre son calculadas aplicando colecciones, de esta manera, la teoría matemática de la probabilidad sólo trata con colecciones bien definidas. Esta restricción de sólo referirnos a colecciones es vista como una limitación, pues sólo trata con eventos cuantitativos, pero recordemos que Russell no sólo se refiere a eventos cuantitativos sino que también son cualitativos que afectan nuestro grado de credibilidad.

Por ello, surgen posturas como la de Russell que exponen los grados de credibilidad que hay en una persona para interpretar los resultados de la teoría matemática de la probabilidad. Pero cuando queremos afirmar el grado de credibilidad, es necesario que nuestra colección básica deba consistir en proposiciones las cuales son igualmente creíbles en relación con la evidencia; en efecto yo no voy a creer en la $P(1/6)$, sino la interpretación que doy a esa probabilidad a ese resultado matemático, que es $P(1/6)$; la interpretación que doy es: creo que caerá el número 6 del dado a lanzar.

Es importante saber qué debemos descartar sobre las características de un evento x para calcular su probabilidad. Por ejemplo, si una bolsa contiene bolas de dos distintos colores, (unas son blancas y otras son negras) en este caso no me es relevante saber las moléculas con que están compuestas las bolas, pero un físico podría objetar y decir que si las bolas están constituidas de distintos materiales, digamos las bolas blancas son menos pesadas que las negras, entonces sí es relevante este dato, pues las bolas ya no estarán distribuidas de forma azarosa, sino que las blancas serán las que estén arriba de las negras y las negras las que estén al fondo de la bolsa.

En la ciencia, como dice Wesley Salmon: "Saber cuándo un dato es considerado como evidencia relevante […] depende del área de investigación y la experiencia les enseña a los científicos a decidir qué dato es relevante"[50]. Ésta es otra respuesta, pero aún no queda del todo claro cuál es el criterio de los científicos para decidir qué debe ser considerado como dato y después qué debe ser considerado como dato relevante.

Los datos son, como ya mencionamos, importantes para Russell, pues de ellos nos basamos para dar probabilidades y los datos que demos pueden cambiar de forma notable la probabilidad. Ahora Russell, define *dato* como una proposición la cual tiene un grado de credibilidad racional en su explicación; las creencias racionales a las que se refiere Russell son las creencias fundadas en las razones que tengamos para tener cierto grado de confianza en un evento, o como él lo llama: grado de creencia racional que, como dice Russell puede tener tres distintos orígenes.

> […] cualquier proposición que es racionalmente creíble en cualquier grado debe ser ya sea:
> 1.-solamente cuando es apreciada ó
> 2.-solamente como la conclusión desde las premisas las cuales son racionalmente creíbles cuando son apreciadas ó
> 3.- porque tiene algún grado de credibilidad que es visto y también se sigue por una demostración ó una inferencia probable, desde las premisas que tienen algún grado de credibilidad. [Russell 1992, p. 409]

Russell comienza a exponer las objeciones o debilidad que tienen nuestras creencias que obtenemos de los tres puntos mencionados. Comienza con objetar la percepción, argumenta que es débil: si estamos escuchando el sonido de algo no sabemos con exactitud cuándo dejamos de oírlo o cuándo comenzamos a oírlo. Ahora bien, la memoria no la considera

---

50 [Salmon 1973, p.85]

como un dato, pues es débil como premisa: si yo quiero saber a qué hora cené con un amigo el año pasado y tengo mala memoria, puedo dar una hora equivocada, pues no confío en mi memoria. Por ello, el hecho pasado que queríamos fundamentar con otras proposiciones no es considerado como dato.

La última objeción sería que desconocemos la conexión lógica: para los que tienen habilidades en lógica es más fácil saber si un argumento es válido, en cambio para un principiante en lógica no lo es, por ello, para el principiante aplicar el método de reducción al absurdo a un argumento no sería un dato concluyente, pues no sabe si lo aplicó bien o mal.

En conclusión, si la lógica deductiva no nos es suficiente para tratar problemas de la vida diaria y de la ciencia, la lógica inductiva tampoco nos es suficiente pero es un comienzo para tratar casos que la lógica deductiva no puede. Ahora bien, la teoría matemática de la probabilidad es más eficiente, pero sólo cuando es aplicable, es decir, cuando los datos a los que se aplica la probabilidad son formales y pueden ser numerables. Sin embargo, en el campo de la ciencia donde se toman decisiones sobre eventos con distintas características, estos eventos, por lo general, tienen características tanto cualitativas como cuantitativas, por ello, se necesita un método que trate estos eventos, es decir, un método que no se restrinja a lo cuantitativo. Russell comienza por proponer la importancia de los grados de credibilidad, que hacen más útil a la probabilidad en problemas que constantemente tratamos, esta propuesta sugiere un método de la probabilidad que trate con eventos que son tanto cualitativos como cuantitativos; el método bayesiano es el que puede tratar estos dos tipos de eventos; además da cuenta de que la probabilidad no se limita sólo a eventos cuantitativos[51]; sin embargo, aún hay muchos factores y circunstancias que debemos tratar, además de saber cómo tratarlos para obtener las hipótesis más probables y elegir la que más nos convenga. Además de solucionar los problemas que hasta ahora han surgido con las propuestas que tratan problemas de la ciencia y la vida diaria. Por último, Russell cuando introduce la teoría de los grados de credibilidad en el cálculo matemático lo hace con la intención de proponer una manera de exponer los limites que tiene la teoría matemática de la probabilidad y cómo esta limitación puede afectar en los procesos de inferencias científicas o de la vida diaria, pues cuando se trata de corroborar una hipótesis, por lo general, nos basamos en los resultados de algún método de probabilidad aplicado a la hipótesis dada la evidencia, sin embargo no sólo debemos tomar en cuenta el resultado del cálculo de probabilidad sino también nuestro grado de credibilidad que está compuesto de nuestras creencias personales o colectivas.

Más adelante trataremos más a fondo las objeciones que se le dan a la probabilidad cuando ésta es utilizada para justificar la selección de hipótesis científicas, comenzando por preguntar ¿cuál es el criterio para preferir una hipótesis sobre las demás? Ahora continuaremos con la teoría subjetiva de la probabilidad.

2.2 Teoría subjetiva de la probabilidad y el teorema de Bayes

Introducción

En el anterior subcapítulo hemos visto como Russell apunta a varios problemas en las inferencias de la ciencia y hace referencia a la atribución de la probabilidad cuando se

---

[51] Podemos notar que el método bayesiano de la probabilidad es más eficiente, pues el método bayesiano puede tratar la postura de Russell sobre los grados de credibilidad, esto lo puede hacer porque, trata con eventos con características cualitativas o cuantitativas.

aplica para la corroboración o refutación de hipótesis[52]. También notamos que está a favor de la probabilidad que puede ser relacionada con grados de credibilidad; esta probabilidad se identifica por lo general con la teoría subjetiva de la probabilidad. También apunta a la importancia que tiene la utilidad que se le da a la probabilidad, pues con base en el uso de la probabilidad podemos notar qué postura filosófica se toma sobre ella.

Luego, en capítulos anteriores hemos explicado de manera rigurosa y formal el teorema de Bayes, ahora explicaremos de forma más histórica y filosófica la teoría bayesiana de la probabilidad, para ello necesitamos explicar la teoría subjetiva de la probabilidad de de Finetti y Ramsey.

Como ya habíamos mencionado, en la actualidad se estudian distintas teorías de la probabilidad desde distintas posturas; las más conocidas son la subjetiva y objetiva, o la distinción cualitativa y cuantitativa (numéricas y no numéricas). La distinción subjetiva y objetiva puede ser entendida como problemas no bayesianos y problemas bayesianos respectivamente, es decir, la distinción se fundamenta en el teorema de Bayes. Históricamente la teoría subjetiva fue originada por Ramsey y de Finetti. La propuesta más radical fue la expuesta por de Finetti que rechazaba nociones objetivas de la probabilidad; más adelante veremos su interpretación subjetiva de la probabilidad.

La teoría subjetiva de la probabilidad

---

[52] Anteriormente se mencionó que, en la introducción del capítulo V del texto de Russell cuando habla sobre 'probabilidad' Russell afirma que hablará sobre las inferencias científicas, pero para ello debe hablar sobre los métodos científicos y hacer una investigación sobre las diferentes clases de probabilidad. Como dice Russell: "This state of affairs makes it evident that we cannot understand scientific method without a previous investigation of different kinds of probability" [Russell 1992, p.354]

La teoría subjetiva de la probabilidad es atribuida a Bruno de Finetti, en Italia, y Frank Ramsey, en Cambridge, aunque los descubrimientos fueron independientes al parecer sucedieron al mismo tiempo, es decir, se cree que de Finetti no conocía el trabajo de Ramsey, y Ramsey tampoco conocía el trabajo de de Finetti, pero ambos parten de un conjunto común de ideas con algunas diferencias.

Antes de la teoría subjetiva, a mediados de 1920 había muchos problemas relevantes en la tradición del bayesianismo lógico (la interpretación de la regla de Bayes bajo la teoría lógica de la probabilidad de Keynes que veremos enseguida). Algunos estadísticos como Fisher y filósofos de la ciencia como Popper miraban con recelo al bayesianismo; sin embargo, la nueva versión del bayesianismo de de Finetti y Ramsey introdujo nuevos conceptos para solucionar los problemas que tenía la interpretación lógica de la regla de Bayes.

Por un lado, según Galavotti[53] Ramsey hizo considerables aportaciones a la filosofía de la probabilidad, a la lógica matemática y a la economía; por otro lado, en 1928 de Finetti expuso ampliamente los fundamentos de la teoría de la probabilidad desde su punto de vista subjetivo.

La teoría de Ramsey surge como una crítica al trabajo sobre la teoría lógica de la probabilidad de Keynes. Para Keynes la probabilidad habla de grados de creencia racional:

> … en el sentido importante para la lógica, la probabilidad no es subjetiva. No está, es decir, sujeta al capricho del humano. Una proposición no es probable porque pensemos que lo es. Cuando algunos de los hechos se dan, los cuales determinan nuestro conocimiento, cuál es probable o improbable en tales circunstancias ha sido fijado objetivamente, y es independiente de nuestra opinión. La Teoría de la Probabilidad es lógica, por lo tanto, porque le concierne con el grado de creencia el cual es *racional* para

53 [Galavotti 2005, p. 48]

> tratar en condiciones dadas y no meramente con la creencia real de individuos particulares, la cual puede o no puede ser racional. [Keynes en Gillies 2000, p.33.][54]

Cuando Keynes habla de probabilidades se refiere a una probabilidad independiente de quien la calcula y para él hay una intuición lógica implícita en todos aquellos que calculan la probabilidad de un evento *x,* pero, también considera más realista una aproximación cualitativa de la probabilidad, más que una numérica.

Ahora bien, Keynes explica los casos donde se puede dar un valor numérico y de este modo asignar probabilidades cuantitativamente, sobre estos casos, él dice:

> En donde la medida numérica es posible debemos dar un número de alternativas igualmente probables (…) una medida numérica puede actualmente obtenerse en estos casos sólo en la reducción de un conjunto de alternativas exhaustivas y exclusivas, pueden ser equiprobables. [Keynes en Gillies 2000, p.35]

Para justificar que los casos son igualmente probables, utiliza el principio de indiferencia o el principio de razones insuficientes. Pero, el principio de indiferencia fue varias veces criticado y en la actualidad sigue siendo criticado, pues su uso origina muchos problemas que ya expusimos en el primer capítulo estos son: la idea de 'equiposibilidad' o 'equiprobabilidad' y la idea de suponer que todos los casos son posibles.

Las críticas que van dirigidas a la teoría lógica de la probabilidad, no sólo son por el uso del principio de indiferencia, también van dirigidas a nociones ambiguas que Keynes introduce en su interpretación de la regla de Bayes. Por ejemplo, según Gillies, Ramsey

---

[54] Este es el texto original de mi traducción: … in the sense important to logic, probability is not subjective. It is not, that is to say, subject to human caprice. A proposition is not probable because we think it so. When once the facts are given which determine our Knowledge, what is probable or improbable in these circumstances has been fixed objectively, and is independent of our opinion. The Theory of Probability is logical, therefore, because it is concerned with the degree of belief which it is *rational* to entertain in given conditions, and not merely with the actual beliefs of particular individuals, which may or may not be rational. [Keynes en Gillies 2000, p.33]

critica la noción de ‘intuición lógica’ que introduce Keynes, pues su explicación sobre esta noción es ambigua, no queda del todo determinada.

Ramsey objeta a Keynes su justificación sobre los grados de creencia racional, pues en su interpretación de la regla de Bayes en la teoría lógica de la probabilidad afirma que el grado de creencia racional es el mismo para todos los individuos racionales, porque si todos partimos de la misma evidencia, supone que todos llegaremos a la misma hipótesis, pues todos tenemos intuiciones lógicas. Sin embargo, partir de la misma evidencia y ser personas racionales no nos garantiza que llegaremos a la misma conclusión, además su noción de ‘intuición lógica’ es criticada, pues es vaga y no está justificada por los axiomas de la probabilidad. Recordemos que Russell también habla de la intuición y de las relaciones lógicas que hay en los argumentos. Para Russell el punto es que las relaciones lógicas simples podemos verlas con facilidad, pero en casos complicados puede ser que mi intuición no esté educada para ver a primera vista estas relaciones lógicas, sino que necesito tiempo para analizar; de este modo la intuición lógica no es la misma. Por ello, las relaciones lógicas que hay en la probabilidad pueden darnos distintas hipótesis, esta postura de Russell también puede ser una crítica a la teoría de la probabilidad de Keynes.

La interpretación subjetiva de la probabilidad abandona el supuesto de que se llega a un consenso racional como afirma Keynes. De modo que la teoría subjetiva define la probabilidad como el grado de creencia de un individuo en particular. Como habíamos dicho, la teoría subjetiva es caracterizada por contrastar la probabilidad con el grado de creencia de un individuo en particular. No debe ser confundida con la teoría lógica de la probabilidad que afirma que todo hombre racional tiene el mismo grado de creencia racional, si parten de la misma evidencia.

La teoría subjetiva abandona la idea de que diferentes individuos, todos racionales y teniendo la misma evidencia *e*, pueden tener iguales grados de creencia en alguna hipótesis *x*. En cambio, considera que cada individuo en específico puede tener distintos grados de creencia, independientemente si parte o no de la misma evidencia que otros.

Cómo medir los grados de creencia

El problema que enfrenta la interpretación de Keynes sobre la teoría lógica de la probabilidad y la teoría subjetiva de la probabilidad es el de cómo medir los grados de creencia. Encontrar una solución fue todo un proceso; primero se intentó medir con un barómetro las sensaciones que tenía una persona cuando ocurría el evento, pero esto es muy ambiguo, pues se pueden tener distintos criterios para proponer las medidas para medir las sensaciones, según Gillies: "Ramsey tiene una interesante discusión sobre este problema. Su primera afirmación sobre la cuestión es que es, supongo, concebible que los grados de creencia puedan ser medidos por un psicogalvanómetro o algún otro instrumento"[55]. Luego de varios intentos, Ramsey concluyó que el grado de una creencia es una propiedad causal, la cual podemos decir que es el grado que nos prepara para actuar de cierta manera. De ahí que debamos medir la fuerza de una creencia, de un individuo, examinando el carácter de alguna de sus acciones que su creencia lo lleva a ejecutar. De esta manera, surge la idea de medir la creencia por medio de la 'apuesta'.

El cálculo de la apuesta

El cálculo de la apuesta es un proceso para medir los grados de creencia de un individuo en específico. Mediante este proceso se dice que se puede medir el grado de creencia por

---

55 [Gillies 2000, p.53]

medio de una situación de apuesta; por ejemplo, se miden los grados de creencia de un individuo *x* sabiendo cuánto está dispuesto a apostar en una situación de apuesta; si el individuo apuesta una gran cantidad de dinero, entonces se considera que su grado de credibilidad es alto, pero si apuesta una mínima cantidad de dinero, entonces se dice que su grado de credibilidad es muy bajo. Según la condición de 'coherencia', un alto grado de credibilidad le corresponde a una gran cantidad de dinero que apuesta el individuo, se espera que el individuo sea coherente, es decir, que él actúe de una forma coherente; su objetivo es obtener un beneficio al ganar una apuesta. Sin embargo, la noción de coherencia puede ser interpretada de distintas formas y por ende se puede medir de distintas maneras los grados de credibilidad de un individuo; de esta forma, obtener distintas probabilidades. Si este es el caso, entonces el cálculo de apuesta no es un criterio seguro para medir los grados de credibilidad.

Ahora bien, hay quienes no están de acuerdo en que se pueda lograr hacer coincidir el cálculo algebraico con los grados de creencia, como Salmon, "puesto que los grados de creencia no coinciden con el valor calculado de la probabilidad de una hipótesis, la interpretación subjetiva no constituye una interpretación admisible del cálculo de probabilidad."[56] Salmon dice que la teoría subjetiva falla en el criterio de admisibilidad que él propone para considerar una interpretación de la probabilidad.

> Admisibilidad: podemos decir que una interpretación de un sistema formal es admisible si los significados asignados a los términos primitivos transforman los axiomas formales y en consecuencia todos los teoremas en afirmaciones verdaderas. Es un requisito fundamental para los conceptos de la probabilidad satisfacer las relaciones matemáticas específicas para el cálculo de la probabilidad. [Salmon 1984, p. 63]

---

56 [Salmon 1984, p.68]

Volviendo a la objeción de Salmon, considera que la interpretación subjetiva de la probabilidad sólo habla de la probabilidad como si fuera una forma de medir los grados de creencia, a mi parecer creo que identifica la interpretación subjetiva de la probabilidad con el principio de indiferencia, pues pone como ejemplo la creencia que tengo de que mañana saldrá el sol. Si yo creo con total certidumbre de que saldrá el sol, mi grado de creencia, según Salmon, es 1, por otra parte si mi grado de creencia de que mañana se acabará el mundo es muy débil de manera que estoy totalmente convencida; según Salmon, mi grado de creencia sobre este respecto es 0. Otro ejemplo es el juego de los volados, si estoy fuertemente convencida de que al lanzar una moneda saldrá cara o cruz, mi grado de creencia y probabilidad subjetiva para cada uno de estos resultados es ½. En esto se basa Salmon para decir que el principio de indiferencia es parecido al de la probabilidad ya que, cuando no tengo evidencia de que el sol no saldrá mañana, es decir, no tengo evidencia de que por lo menos una vez no haya salido el sol, y en el lanzamiento de los volados no tengo evidencia de que la moneda pueda estar trucada o no, por falta de esta evidencia mi grado de creencia seguirá siendo el que siempre he tenido (1 para el caso de que saldrá el sol mañana y ½ para el caso de que es posible de que salga cara o cruz en un volado). Por ello, mi grado de creencia no se verá afectado por ningún evento, ya que no los hay, por consiguiente, no tengo razones para elegir una hipótesis sobre otra, en eso, según Salmon, se parece a la interpretación subjetiva de la probabilidad.

Esta es la postura de Salmon sobre la interpretación subjetiva de la probabilidad, por eso considera que es gravísima la adecuación de un concepto puramente subjetivo dentro de la aplicación de la ciencia empírica. Para él la admisibilidad de una interpretación del cálculo de probabilidad consiste en interpretar los axiomas y los teoremas de tal manera que los resultados correspondan o sean coherentes los resultados con los axiomas, es decir, que los resultados no sean afectados por una creencia subjetiva, pues si es así, podemos obtener varios resultados que serán válidos todos pero perderíamos rigor por ello.

Sin embargo, pronto veremos que la teoría subjetiva de la probabilidad de de Finetti no interpreta la aplicación de los grados de creencia sobre la probabilidad de la forma en que Salmon cree, pues por ejemplo, en el juego de los volados, de Finetti considera una secuencia de volados para calcular la probabilidad de que en el próximo lanzamiento saldrá cruz o cara; no se queda en la interpretación de la probabilidad de que será ½ como lo hace el principio de indiferencia de Laplace. Pero volviendo a las objeciones de Salmon, si aceptamos la interpretación subjetiva tal como Salmon la expone, podemos ver que no puede satisfacer el criterio de admisibilidad. Salmon expone el siguiente ejemplo:

> Los jugadores del siglo XVII, por ejemplo, creían en un grado de 1/36 en cada uno de los posibles resultados del lanzamiento de dos dados, y este grado no fue afectado por los resultados de los últimos lanzamientos. Además estos mismos hombres creían en obtener al menos un doble seis en 24 lanzamientos. […] El cálculo mostró, como resultado, que la probabilidad de conseguir doble seis, en las condiciones especificadas, es inferior a la mitad. [Salmon 1984, p. 68]

De esta forma Salmon justifica el por qué cree que los grados de creencia no pueden coincidir con los valores calculados, por ello, la interpretación subjetiva no constituye una interpretación admisible del cálculo de probabilidad. En cambio, para de Finetti satisfacer los axiomas de la teoría de la probabilidad utilizando conceptos como 'coherencia' es necesario y suficiente para considerar la teoría subjetiva de la probabilidad como admisible. Esta exposición se encuentra en el texto de Gillies[57]. Aunque Salmon distingue la probabilidad a priori y a posteriori, reitera que la probabilidad a priori no es calculada de manera objetiva, es decir, para él la probabilidad a priori sólo está fundada en creencias

---

[57] [Gillies 2000, p.58-65]

injustificadas de manera rigurosa y esta falta de justificación afecta el cálculo de probabilidad de manera que pierde rigor el método bayesiano.

Por último, Gillies no está del todo convencido de que la apuesta sea una manera razonable de medir la fuerza de una creencia, pues cree que hay casos en que no se aplica; su postura es que la apuesta da una medida razonable de la fuerza de una creencia en algunos casos, pero no en todos. Acepta, por el momento, que apostar es una postura razonable de medir el grado de creencia.

La noción de 'coherencia'

El problema de cómo medir los grados de credibilidad parece solucionarse, pero ahora el siguiente problema es justificar que un conjunto de cocientes de apuestas es 'coherente'. El conjunto de apuestas es coherente para de Finetti si y sólo si satisface los axiomas de probabilidad, esto es, la teoría subjetiva se debe probar por medio de un cálculo en donde se debe obtener como resultado que los grados de creencia llegan a satisfacer los axiomas fundamentales de la probabilidad. Es un requisito fundamental que debe cumplir cualquier teoría que se quiera postular como una teoría rigurosa y formal de la probabilidad.

La solución a la derivación de los axiomas es por medio del uso del concepto de coherencia; se demuestra formalmente que el uso de este concepto es tanto necesario como suficiente para satisfacer los axiomas de la probabilidad.[58]

---

[58] Por cuestiones de seguimiento al tema de este texto el cual no es dar una prueba formal de los métodos de probabilidad y para no entrar en tecnicismos matemáticos sólo mencionare el texto en donde se puede encontrar la prueba matemática para satisfacer los axiomas, puede consultar el siguiente texto [Gillies 2000, pp. 59-64]. En la prueba se utilizó el teorema de de Finetti - Ramsey. El teorema muestra satisfacer los axiomas de la probabilidad, lo que no se pudo hacer en la teoría lógica de la probabilidad, pues sólo fueron justificados los axiomas por un uso insatisfactorio y ambiguo de la noción de 'intuición lógica' que Keynes introduce.

En la teoría subjetiva se puede demostrar formalmente apelando a la condición de coherencia, ya que los axiomas de la probabilidad son satisfechos cuando se introduce la condición de coherencia y se aplica a la teoría subjetiva de la probabilidad de de Finetti. Sin embargo, la noción de coherencia a su vez tiene varias interpretaciones, todas pueden mostrar ser rigurosas, por ello, podemos tener varias interpretaciones de cómo medir la coherencia de las creencias de un individuo sobre un evento particular *x*. Por ejemplo:

> Tomoji Shogenji (1999) defiende que la medida de la coherencia es igual a la relación del conjunto de probabilidad de todas las proposiciones en *S* sobre el producto de las probabilidades marginales de las proposiciones en *S*. Erik J. Olsson (2002b: 250) sugiere como una posible medida de la coherencia la relación de la probabilidad conjunta de las proposiciones en *S* sobre la probabilidad de la disyunción de las proposiciones en *S*. Branden Fitelson (2003) defiende una medida de la coherencia que está basada sobre la medida de Kemeny y Pals del apoyo fáctico. [Bovens 2003, p.12]

De esta forma, podemos ver que no hay un criterio seguro por medio del cual podamos decir que tal interpretación de la medida de coherencia es mejor que otra.

Cálculo de eventos que parecen objetivos

Otra dificultad que enfrenta la teoría subjetiva es cómo calcular la probabilidad de eventos que parecen objetivos, (para de Finetti todos los eventos empíricos que trata la ciencia como evidencia son subjetivos y sólo parecen ser objetivos); si bien la probabilidad subjetiva se inclina por calcular la probabilidad de eventos relacionando grados de creencia de individuos en específico, es decir, hay probabilidades que envuelven un componente de la subjetividad: la probabilidad de que mañana llueva, de que gane tal partido las elecciones. Sin embargo, si hay otras probabilidades que parecen objetivas, ¿cómo calcular la probabilidad de eventos que parecen ser de otro concepto distinto a la interpretación subjetiva de la probabilidad? Es decir, eventos que son independientes del grado de creencia del individuo y que parece que se presentan en diferentes circunstancias.

De Finetti trata los eventos que parecen ser objetivos introduciendo el concepto de intercambiabilidad; afirma que “se puede decir que todas las probabilidades son subjetivas y que incluso son aparentemente objetivas, tal como puede ser explicado en términos de grados de creencia subjetiva”[59]. De Finetti reconoce que hay tal tipo de eventos y los llama *eventos intercambiables*.

Ahora bien, si nos parece que hay distintos conceptos de probabilidad, también nos puede parecer que sólo hay uno, como dice de Finetti se puede decir que todas las probabilidades son subjetivas; creo que este es un problema esencial en la noción de probabilidad que adoptemos, pues no hay una noción universal de probabilidad y al carecer de ella podemos creer que cualquier noción de probabilidad no es mejor que la nuestra; por ejemplo: la postura subjetiva de de Finetti considera que todas las probabilidades se pueden interpretar desde su teoría de la probabilidad, pero también Gillies puede decir que todas las probabilidades se pueden interpretar desde la teoría objetiva de la probabilidad. Independientemente de qué teoría de probabilidad adoptemos debemos estar conscientes de los problemas, las dificultades que podría tener y de las ventajas que obtenemos si la utilizamos, como Gillies dice al principio de este tema, él considera que la teoría subjetiva tiene un mejor uso en las ciencias sociales y la teoría objetiva en el campo científico.

Sin embargo, hay otra dificultad cuando se trata de saber qué teoría de la probabilidad es mejor (más eficaz), pues como veremos en el tema: el problema de la

---

59 [de Finetti en Gillies 2000, p.69]

interpretación de la evidencia, Kuhn y su concepto de 'inconmensurabilidad', es difícil decir qué teoría de la probabilidad es mejor, pues cada una tiene su forma de ver el mundo, cada una clasifica los eventos de distinta forma, y cuando se quieren comparar debería de tomarse en cuenta esta diferencia, para, por lo menos, ser conscientes de la dificultad de querer comparar una teoría de la probabilidad con otra. Por ejemplo, ya vimos una clasificación de eventos numéricos y no numéricos en el capítulo 1, esta interpretación que da la teoría frecuencialista de la probabilidad tiene el problema de decidir qué eventos deben ser considerados como numéricos y cuáles no. Otro tipo de interpretación es la frecuencialista que según Gillies clasifica los eventos como aquellos que tienen una secuencia ordenada y aquellos que no tienen una secuencia ordenada; pone como ejemplo de una secuencia ordenada los volados y como un evento no ordenado, pone a la secuencia de contar flores en un campo, no tenemos un orden; si ésta es su forma de clasificar tiene el problema de justificar cuáles eventos deben ser considerados como ordenados y cuáles no, es decir, debe de tener una buena justificación o un buen criterio. De aquí que tenga relación con el problema filosófico de la percepción que Hanson[60] nos muestra: varias perspectivas, distintos puntos de vista que tratan de estar de acuerdo en algo, pues para algunos una secuencia ordenada tendrá ciertas características, pero otra persona puede creer que ciertas características no son suficientes y/o necesarias para definir una secuencia ordenada. Aunque en la probabilidad lo único en común que deben de tener las teorías formales y rigurosas es que deben satisfacer los axiomas de la probabilidad, pero no es suficiente para saber qué probabilidad es mejor aplicar.

---

[60] Más adelante en el tema: el problema de la observación de la evidencia y sus efectos en los métodos científicos, veremos el problema de la percepción a la que Hanson se refiere. Este problema trata sobre la observación de un evento, pues habrá ocasiones en que haya discrepancia de opiniones al respecto del mismo objeto observado.

Volviendo al problema inicial, si hubiera casos objetivos, la solución que expone de Finetti es introducir la condicionalización de intercambiabilidad y considerar que pueden ser explicados en términos de grados de creencia, pero antes debemos explicar la condicionalización bayesiana de de Finetti. El siguiente ejemplo puede ayudarnos a entenderlo. Supongamos que en el juego de los volados se quiere calcular la probabilidad de obtener cara en el siguiente volado, si contamos con una secuencia de lanzamientos de la moneda que da como resultados: $E_1,\ldots, E_n,\ldots$, donde cada E_i puede ser o cara (H_i) o cruz (T_i). Ahora bien, si denotamos H_{n+1} como las caras ocurridas sobre los n+1 volados, que es una secuencia de n lugares, en el 'i' lugar de los cuales tenemos ya sea cara (H_i) o cruz (T_i). Supongamos que cara ocurrió 'r' veces sobre los primeros n volados. Si varios individuos saben esto, pueden tener una probabilidad *a priori* de que suceda cara en el siguiente volado y de cuánto sea esta probabilidad *a priori* depende de ellos, es decir de qué tan confiados estén con respecto a la información que tienen. Después de tener sus probabilidades *a priori*, el método subjetivista es calcular $P(H_{n+1}\ \mathrm{I}e)$ donde e = evidencia, y para mostrar que sobre algunas condiciones generales las cuales son especificadas después $P(H_{n+1}\ \mathrm{I}e)$ tiende a r/n de una larga serie de eventos n.

Esto muestra que el valor es asignado como la probabilidad *a priori* $P(H_{n+1})$, y la probabilidad *a posteriori* $P(H_{n+1}\ \mathrm{I}e)$ tenderá a la frecuencia observada de una larga secuencia de n eventos.

De esta manera de Finetti quiere demostrar que diferentes individuos pueden sostener diferentes opiniones inicialmente y si ellos cambian su probabilidad por la condicionalización bayesiana, $P(H_{n+1}\mathrm{I}e)$, llegan a un acuerdo sobre sus probabilidades *a* *priori*. Luego se fundamenta en esto para decir que la probabilidad objetiva es un concepto metafísico desprovisto de significado, pues considera que cuando se llega a un acuerdo sobre distintas probabilidades *a priori*, para nosotros nos parecerá que es una probabilidad *a posteriori*, pero para de Finetti no, pues considera que sigue siendo *a priori*, se puede creer que cuando se llega a este acuerdo se es objetivo más que subjetivo, pero de Finetti defiende que no es así, pues sólo se puede llegar a un acuerdo cuando se está en desacuerdo sobre las probabilidades *a priori*. Según él, todo sucede, a la luz de evidencia, diferentes individuos llegan a un acuerdo sobre sus probabilidades subjetivas. Por eso, de Finetti cree que la probabilidad objetiva supone una ilusión de que hay una real probabilidad objetiva p. Diferentes personas pueden, relacionarse sólo por la coherencia, pero tenemos diferentes probabilidades *a priori* $P(H_{n+1})$.

$$P(H_{n+1} \mid e) = \frac{H_{n+1} \,\&\, e}{P(e)} \qquad [1]$$

Ahora bien, para calcular la probabilidad de caras r veces en n volados, de Finetti introduce la condición de intercambiabilidad como, por ejemplo: supongamos que un individuo en específico digamos el señor B está calculando una hipótesis *a priori* sobre un particular n de resultados (E_{i1}, E_{i2}, …, E_{in}) de dichas ocurrencias. Supongamos que en un futuro ocurrirá caras r veces en n volados y el señor B hace un cálculo de apuesta para decir que en un futuro, cierto evento podría suceder, pero si hace varios cálculos de varios posibles eventos que podrían suceder, entonces tendrá varias hipótesis *a priori* y estos *a priori* serán intercambiables si tiene la misma probabilidad después de haber sido calculados con el cálculo de apuesta ya mencionado.

De este modo, de Finetti utiliza la condición de intercambiabilidad para tratar con eventos, que en términos objetivistas se llamarían eventos independientes, pues se refieren a los mismos eventos y tienen la misma probabilidad cada resultado.

En resumen, "coherencia + intercambiabilidad + algunos otros supuestos plausibles [...] asegura que $P(H_{n+1}$ I$e) \rightarrow r/n$ como $n \rightarrow \infty$."[61] Así es como la evidencia se acumula y las personas que no están de acuerdo en una probabilidad *a priori* en un futuro llegarán a un convenio o estarán de acuerdo en una probabilidad *a posteriori*. Ahora bien, de Finetti dice que la probabilidad objetiva es una gran ilusión, pues, del hecho de que se pueda llegar a un acuerdo *a posteriori*, parece que da la ilusión de una probabilidad objetiva, pero en realidad la probabilidad *a posteriori* está fundamentada en probabilidades *a priori* (subjetivas). Como habíamos dicho en el primer capítulo el concepto de *a priori* y *a posteriori* son vistos como la probabilidad *a priori* de una hipótesis, es decir, es el conocimiento que antecede a la hipótesis y es relevante para calcular su probabilidad; luego, la probabilidad *a posteriori* es la probabilidad que obtenemos después de calcular la probabilidad de la hipótesis considerando su probabilidad *a priori*[62].

La teoría subjetiva parece solucionar la paradoja del principio de indiferencia para hacerlo innecesario. En cambio en la teoría lógica de la probabilidad, el principio de indiferencia[63] se usó para obtener supuestamente el único grado de creencia racional *a priori.* Sin embargo, la teoría subjetiva no considera que hay sólo una probabilidad *a priori*.

---

[61] [Gillies 2000, p.72]

[62] Puede consultarse la explicación de la noción *a priori,* de una forma más detallada, en la p.98 de este texto.

[63] El principio de indiferencia tiene varias objeciones que han sido desarrolladas, si el lector desea saber más sobre estas objeciones y posibles soluciones puede consultar en el texto de Gillies 2000, pp. 37-49.

Si bien, los diferentes individuos pueden elegir su propia probabilidad *a priori* en diferentes posturas y probar que son coherentes (aquí de nuevo aparece el problema de la coherencia, pues no hay una definición única de este concepto), entonces no parece haber problemas con estas diferentes opciones. Sin embargo, entre los problemas que aparecen conectados a la teoría subjetiva, está el problema de la noción de intercambiabilidad que de Finetti introduce en su teoría subjetiva de la probabilidad para no usar términos objetivos que cumplen la misma función que de intercambiabilidad; por ejemplo: la noción de intercambiabilidad es equivalente a la noción de independencia, es decir, cumple la misma función que la noción objetiva de independencia o equiprobabilidad. La noción de intercambiabilidad expresa que la probabilidad de A no depende de B, esto es simbolizado:

$P\ (AIB) = P(A)$ y si la $P\ (BIA) = P\ (B)$, entonces A, B son mutuamente independientes.

Esta definición en el proceso, puede ser aplicada cuando las probabilidades se relacionan con un significado de subjetividad. El problema surge cuando en la aproximación objetiva se aplica en muchos casos, la independencia en el sentido subjetivo (intercambiabilidad) no hay un buen criterio para considerar qué eventos son intercambiables y cuáles no, pero ésta es una crítica que veremos en el tema de deficiencias del método bayesiano.

De Finetti llama a las probabilidades objetivas como sólo apariencias, porque para él sólo parecen objetivas cuando a la luz de nueva evidencia, diferentes individuos llegan a estar de acuerdo en sus asignaciones subjetivas de probabilidad. Si bien se criticaba la teoría subjetiva de de Finetti por considerar varias probabilidades *a priori*, o hipótesis iniciales (que tan bien pueden ser consideradas como probabilidades *a priori* que se tiene sobre cierta hipótesis) a la luz de la evidencia obtenida en los primeros experimentos o fenómenos que son observados, se critica cómo es posible que lleguen a un acuerdo, es decir, hay varias probabilidades *a priori*, de éstas ¿cómo se elige la mejor? De Finetti responde que se llega a un acuerdo cuando los individuos que tienen distintas probabilidades y después obtienen más evidencia, llegan a una probabilidad *a posteriori* donde convergen sus opiniones, entre más experimentos u observaciones se tengan, más se acerca a una misma probabilidad *a posteriori*. Parece una respuesta plausible, pero como veremos más adelante Gillies expone varias deficiencias.

2.3 Teoría frecuencialista de la probabilidad

Introducción

Por lo general, una interpretación subjetiva de la probabilidad habla de grados de credibilidad que son tratados con las probabilidades numéricas, estos grados pueden ser de creencia o de duda; la mayoría de estas interpretaciones de la probabilidad se refiere a afirmaciones no numéricas. En cambio, una interpretación objetiva de la probabilidad habla de eventos que tienen la característica de ser numéricos, ordenados y frecuentes.

La teoría objetiva de la probabilidad considera la probabilidad como un cálculo formal de frecuencias de eventos objetivos, por consiguiente, en esta postura se encuentra la teoría frecuencialista de la probabilidad, pues, recordemos que define el resultado de la aplicación de la probabilidad como el límite de la frecuencia donde el evento a considerar aparece en una gran secuencia.

En el capítulo anterior hemos explicado de manera rigurosa y formal la probabilidad frecuencialista, ahora se explicará de una manera más histórica y filosófica bajo la interpretación objetiva de la probabilidad. A continuación, expondré de forma breve los orígenes de la teoría frecuencialista de la probabilidad o también conocida como la teoría objetiva de la probabilidad.

La teoría frecuencialista

La teoría frecuencialista de la probabilidad se desarrolló a mediados del siglo XIX por la escuela de Cambridge. En el siglo XX, destacan como personajes principales de esta teoría dos seguidores del círculo de Viena: Reichenbach y Richard von Mises, ambos coinciden en la misma teoría aunque con algunas diferencias.

La interpretación de von Mises considera que la teoría de la probabilidad frecuencialista trata con fenómenos de materia o masa y eventos repetitivos con un gran número de elementos uniformes de una manera formal, por ello, concluye que la probabilidad es una ciencia matemática. Aunque para algunos es una rama de la lógica, pues, desde la teoría lógica de la probabilidad se considera la probabilidad como una extensión de la lógica deductiva a la lógica inductiva.

Von Mises considera la probabilidad como una ciencia rigurosa y formal, un mecanismo que se relaciona con diferentes fenómenos observables que son repetitivos. En 1919 publicó su primer texto sobre la teoría frecuencialista, después en 1928 publica *Probability, Statistics and Truth;* luego, en 1964 se publica su texto póstumo *Mathematical Theory of Probability and Statistics*. En sus textos se nota la importancia de la noción de colectivo en su interpretación de la probabilidad, pues dice que no se puede hablar de probabilidad si no se tiene un colectivo.

La noción de colectivo trata con eventos objetivos y asocia el resultado de la probabilidad de forma necesaria con colecciones de eventos. Los eventos que menciona von Mises los clasifica en 3 principales grupos: 1) Juegos de azar, 2) Situaciones como las estudiadas por la estadística biológica o por las compañías de seguros, 3) Situaciones que ocurren en la física; por ejemplo, el comportamiento de las moléculas de un gas en específico. Sin importar en qué grupo se encuentren los eventos todos tienen la propiedad de ser repetitivos y observables; sin embargo, von Mises advierte que este atributo varía de un elemento a otro.

Colectivo empírico y colectivo matemático

Von Mises define el *espacio atributivo* Ω o *espacio muestral* Ω como el conjunto de los posibles resultados de experimento aleatorio. Utiliza el término *colectivo* para referirse a eventos no sólo objetivos sino también repetitivos y uniformes, o procesos que se distinguen por ciertos atributos observables: color, número, cualquier cosa que consideremos que la distingue. De estos eventos repetitivos, también hace una distinción entre: colectivos empíricos y colectivos matemáticos. El colectivo empírico se refiere a eventos que existen en el mundo y tienen la característica de ser observados; por ejemplo: volados, moléculas de un frasco, luego se puede inferir que son numéricos y tienen sólo un número finito de eventos si son observables. En cambio, el colectivo matemático lo define como una secuencia ordenada numérica e infinita; por ejemplo, si nosotros queremos calcular la probabilidad de que cierto objeto le corresponda un número real que está entre el 0 y el 1, pues hay infinitos números reales entre el 0 y el 1.

Von Mises da un procedimiento para calcular la probabilidad. Su forma de proceder comienza con localizar los colectivos empíricos, después los representa en la teoría matemática de la probabilidad por medio de un colectivo matemático infinito; el problema que inmediatamente surge es cómo representar un colectivo finito en un colectivo infinito, para von Mises sí es legítima la representación, a pesar de las dificultades que se pudieran presentar, y afirma que la interpretación de lo finito por lo infinito es referida como satisfactoria en cálculos o mecanismos, aunque aún sigue siendo un problema su justificación, pues es imposible tener una secuencia observable y al mismo tiempo infinita. Después, admite que para poder representar el colectivo empírico debe hacer abstracciones matemáticas desde la idealización o realidad empírica; estas abstracciones son necesarias para lograr su representación objetiva y calcular la probabilidad de un evento.

Definición de frecuencia límite de la probabilidad

Si bien la relación entre colectivos empíricos y matemáticos parece no ser isomórfica, la forma en que von Mises relaciona la materia empírica con las ciencias matemáticas es procediendo de la observación a la abstracción del cálculo y, de esta forma, obtener la probabilidad del evento. Pero para proceder de esta forma, Von Mises debe establecer bases de observación a algunas leyes empíricas que el fenómeno debe obedecer; esto elimina hablar de la observación individual y de esta forma podemos evitar que los distintos ordenes que puede tener una secuencia afecte el resultado del cálculo de la probabilidad de cierto evento; al igual que evita los problemas de la observación que Hanson señala y que expondremos en el tema: el problema de la interpretación de la evidencia.

Por consiguiente, otra vez por abstracción o idealización obtenemos desde estas leyes empíricas los axiomas de nuestra teoría matemática para tratar los colectivos empíricos. De esta forma, von Mises justifica la representación de colectivos empíricos en colectivos matemáticos y puede deducir por la lógica consecuencias desde esta representación; además estas predicciones dan explicaciones sobre fenómenos futuros observables y pueden saber la probabilidad del fenómeno estudiado. En el siguiente esquema podemos ver de una forma gráfica la manera de proceder de von Mises.

Observación del fenómeno ______ abstracción o → Concepto matemático

ejemplo: colectivo empírico idealización ejemplo: colectivo matemático

Leyes empíricas ______ abstracción o → Axiomas matemáticos

idealización

↓ deducción

Futuro fenómeno observable ← Explicación o predicción ______ Consecuencias

[Gillies 2000, p.91]

El esquema que presenta Gillies tiene la finalidad de mostrar la visión de von Mises sobre cómo se aplica la teoría de la probabilidad para demostrar que las leyes empíricas obedecen a colectivos empíricos y cómo estas leyes son establecidas. Según von Mises son leyes empíricas de la probabilidad.

La más controversial y conocida de estas leyes es la ley de estabilidad de frecuencias estadísticas. Esta ley establece que la frecuencia de un evento entre más grande sea su frecuencia más se acerca a un valor determinado y se debe aplicar a eventos repetitivos; por ejemplo, en el lanzamiento de dados muestra la frecuencia relativa de ciertos atributos que llegan más y más a establecerse a un valor fijo conforme el número de observaciones aumenta, de esta forma se incrementa la estabilidad de frecuencias de la teoría de la probabilidad; es decir, la frecuencia relativa afirma que cuando se incrementan los sucesos de eventos, se llega a un valor cada vez más fijo, la frecuencia relativa se calcula con la ecuación: $m(A)/n$ donde A es el atributo del evento m y n es la secuencia de todos los resultados.

De esta manera, es esencial que la frecuencia relativa de ciertos atributos tenga más instancias, para obtener una probabilidad más definida. Von Mises se refiere a este incremento de la estabilidad de frecuencias estadísticas como el 'fenómeno primario' de la probabilidad[64].

Volviendo a la ley de estabilidad de frecuencias estadísticas. También puede ser llamada 'axioma de convergencia', este axioma afirma que si A es un atributo arbitrario de un colectivo (c), entonces el límite $_{n\to\infty} m(A)/n$ existe[65].

Reichenbach

Reichenbach afirma que los científicos empíricos, por lo general, se enfrentan a secuencias y tales secuencias son los fenómenos empíricos que estudian. Para él una proposición sólo tiene significado si se le puede asignar una probabilidad; afirma que dos proposiciones tienen el mismo significado si se demuestra que tienen el mismo grado de probabilidad, sin

---

64 Una manera formal de expresar la ley de la estabilidad de frecuencias estadísticas, es: Si A es un atributo arbitrario asociado con un colectivo particular y si Ω es el espacio de atributo (espacio muestral) del colectivo, entonces $A \subseteq \Omega$. Ahora bien, supongamos que en los primeros n miembros del colectivo A ocurre $m(A)$ veces, entonces la frecuencia relativa es $m(A)/n$, es decir, si n incrementa en $m(A)/n$, nos acercamos cada vez más a un valor exacto. Pero esto sólo lo vemos en ejemplos parecidos al juego de los volados, en los naipes y en eventos en donde la secuencia es ordenada y partir de datos establecidos.

65 La probabilidad frecuencialista adquiere un sentido importante con la definición límite de John Venn, esta definición fue expuesta en su texto "*chance logic*" de 1886. Von Mises, Reichenbach entre otros seguidores de la interpretación frecuencialista de la probabilidad se basan en esta definición. Todos ellos afirman la definición límite de Venn, la cual dice: 'si un suceso ocurre un gran número de veces, entonces la probabilidad de ocurrencia del suceso es el límite cuando el número de pruebas tiende a infinito del cociente entre el número de veces en que se presenta el suceso y el número total de pruebas.'

embargo se le puede objetar que podemos suponer que la probabilidad de que una moneda caiga cara es $P(1/2)$ y la probabilidad de que un bebé nazca hombre es $P(1/2)$, pero nos queda claro que la proposiciones respectivas no significan lo mismo. Al igual que von Mises, Reichenbach considera que la experiencia proporciona una base a las expectativas de eventos futuros que nos permite estimar la probabilidad de ocurrencia, tales eventos son miembros de clases y deben ser más o menos repetibles, tener una secuencia; para Reichenbach, aunque se trate de acontecimientos únicos; por ejemplo, cuando un veterinario afirma que un animal enfermo probablemente murió de alergia a la penicilina, el veterinario se basa en su experiencia para estimar dicha probabilidad.

Un elemento central en la epistemología de Reichenbach es su noción de 'postulado'; un postulado es una proposición que es tratada como si fuera cierta, por lo menos temporalmente, aunque no se sabe si lo sea; el postulado sólo se refiere a eventos que poseen la máxima probabilidad y se considera al postulado que es más probable que ocurra, ésta es la forma racional de actuar, y considera que es la más usual, porque es la más práctica. Luego, Reichenbach afirma que la meta de la inducción debe ser una serie de eventos que tienen una secuencia que converge en un límite; la secuencia debe ser de eventos que tienen la cualidad de poder ser observados varias veces. Reichenbach afirma que la intención que tiene cuando describe el procedimiento es sólo hacer una reconstrucción racional del conocimiento científico.

Conclusión.

Por lo que hemos dicho en el capítulo, podemos concluir que la teoría objetiva de la probabilidad no reconoce cualquier evento como susceptible de calcular su probabilidad, sino sólo a aquellos que tienen las características ya señaladas. De esta forma la restricción de sólo aceptar ciertos eventos se ha transmitido a la teoría frecuencialista y de esta manera es un método formal que trata de evitar toda clase de características subjetivas, por ello, tiene la desventaja de dejar fuera varios fenómenos cualitativos que, por lo general se relacionan con la investigación científica como Russell destaca.

Si bien, en el texto de Russell[66], podemos ver que Russell afirma que Reichenbach dice que nuestro conocimiento es sólo probable; Russell afirma que las leyes de la ciencia sólo pueden ser probables, y se pregunta por la justificación de las leyes que infiere la ciencia.

Luego, Russell afirma que cierta clase de hechos particulares son considerados como evidencia de leyes generales. Para Russell, las leyes generales justifican nuestras creencias y cualquier conocimiento que tengamos será considerado como una experiencia particular, si no tiene estas leyes; estas leyes no son las que se construyen a partir de hechos particulares, pero están implícitas en las leyes que formamos a partir de hechos particulares, son leyes que en ocasiones no estamos conscientes de ellas; por ejemplo: si *A* se sigue de *B*, y un animal ve que sucede *A*, él esperará que se siga *B*, supongamos que el animal puede decir ya vi *A*, ahora espero que suceda *B*, sin tener que conocer la ley general. En este caso se refiere a la ley de la causalidad. De estas leyes Russell distingue que tienen cierto estado de probabilidad de la siguiente forma; por ejemplo, es probable que si estudio la estructura lógica que hay en los argumentos, entonces entienda mejor cuando un texto puede contener un argumento.

---

66 [Russell 1992, pp. 300-370]

Es una inferencia deductiva, pero en su contenido implica probabilidad, por ello, la probabilidad es importante, pues, como dice Russell:

> En las premisas de la ciencia no sólo necesitamos datos derivados desde nuestra memoria y percepción, sino también ciertos principios de inferencia científica que obviamente no pueden ser establecidos por la lógica o por argumentos desde la experiencia, puesto que estos son presupuestos en todas las inferencias desde hechos de la experiencia hasta otros hechos o leyes. [Russell 1992, p. 355]

Russell busca principios que justifiquen las leyes de la ciencia y sabe que no es suficiente con las leyes generales, por ello, recurre a los principios de la probabilidad, sin embargo, está el problema de que no hay una probabilidad como 'la probabilidad'; sólo hay teorías, de las cuales adopta una llamándola 'grados de credibilidad'.

Russell atribuye a las premisas de los argumentos de la ciencia grados de incertidumbre que después definirá como grados de credibilidad siguiendo la teoría de Reichenbach.

Pero la anterior teoría de Russell no es la única propuesta para justificar las inferencias de la ciencia; también está la teoría matemática de la probabilidad. En esta definición de la probabilidad sólo se limita a casos donde conocemos un fenómeno porque sucede frecuentemente, es decir, sucede usualmente. Aquí 'probablemente' no tiene el sentido que tiene cuando estamos hablando de grados de probabilidad. Por estas razones Russell expone las dos teorías de la probabilidad que han intentado justificar las inferencias de las ciencias: la que apoya, los grados de credibilidad y la teoría matemática de la probabilidad.

III. El uso de la probabilidad para la corroboración de hipótesis científicas basándose en la evidencia por medio de métodos científicos

Introducción

En este capítulo se expondrán algunos de los problemas filosóficos relacionados con los métodos científicos que se discuten en la filosofía de la ciencia, a saber: la observación de un evento, la justificación de la inducción, la corroboración de hipótesis científicas. Estos problemas que podemos encontrar en el ámbito de la filosofía de la ciencia, creemos que deben ser expuestos, pues de esta forma podemos exponer razones para decir por qué tal método de la probabilidad es mejor a comparación del otro que no trata o no soluciona este tipo de problemas. Por ejemplo, en el problema de la observación de la evidencia, la dificultad que conlleva la observación de la evidencia, puede originar que asumamos distintas interpretaciones de la evidencia y, por ende, distintas razones para apoyar una hipótesis dada tal evidencia. Ahora bien, el método de probabilidad que por lo menos trate este problema o de una solución, puede ser visto como un método de la probabilidad con ventajas en su uso. De los dos métodos de probabilidad que exponemos en este texto, el método bayesiano de la probabilidad trata este tipo de problemas, en cambio el método frecuencialista de la probabilidad no lo hace.

En lo que sigue aceptaré la distinción que Popper[67] afirma entre el contexto de descubrimiento y el contexto de justificación de los descubrimientos científicos[68],

---

67 [Popper 1977, pp. 30-31].

68 Según Popper, el campo de estudio de la filosofía de la ciencia se distingue del campo de estudio del descubrimiento científico que corresponde a las ciencias psicológicas o sociales, ya que, Popper afirma que el contexto de descubrimiento se pregunta por las circunstancias sociales o psicológicas por las cuales el científico originó su descubrimiento; en cambio en filosofía de la ciencia, según Popper, se pregunta por la justificación de tal descubrimiento, es decir, pregunta por las razones por las cuales se afirma tal descubrimiento científico. Popper fue uno de los primeros filósofos en hacer esta distinción en su texto de *la*

ocupándome únicamente del segundo. Por consiguiente, podemos decir que nos limitaremos a tratar las justificaciones y las críticas a los métodos científicos más reconocidos y pertinentes. Luego abordaremos un tema fundamental en filosofía de la ciencia a saber, la observación de la evidencia y el conocimiento que obtenemos a través de la evidencia.

3.1 Corroboración de hipótesis científicas

Salmon (1984) aporta una crítica hacia los argumentos inductivos que son usados en la ciencia para corroborar o invalidar (rechazar) hipótesis científicas[69]. Una hipótesis es corroborada cuando se tiene un caso corroborado de una de las consecuencias de la afirmación de la hipótesis, pero si una consecuencia de la hipótesis es falsa, se dice que se tiene un caso que invalida la hipótesis, estos casos de corroboración o invalidación son reconocidos, si la hipótesis es adecuadamente apoyada por la evidencia inductiva. Aunque Salmon asegura que una hipótesis puede tener grados de corroboración como: altamente corroborada, moderadamente corroborada, ligeramente corroborada. También las hipótesis pueden ser de varios tipos, entre ellos, pueden ser generalizaciones universales que son inferidas de forma deductiva como algunas leyes de la física, Salmon nos da un ejemplo:

---

lógica de la investigación científica, él aclara que el contexto de descubrimiento científico se refiere a las circunstancias sociales; en cambio, el contexto de justificación se refiere a la validez y justificación de las hipótesis científicas, es decir, después de que el científico afirma una hipótesis, la justificación de ésta es lo que puede tratar el filósofo. Tales cuestiones sólo surgen después de que las teorías han sido formuladas. También Salmon en su texto *The Foundations of Scientific Inference.* University of Pittsburg Press. Expone ésta distinción entre justificación y descubrimiento [Salmon 1984, pp.108-110].

[69] Las hipótesis que consideraremos pueden llegar a ser teorías. Vamos a entender por hipótesis científicas su definición más general: son afirmaciones de la naturaleza en general, estas afirmaciones funcionan como hipótesis que se suponen como ciertas, se pueden examinar y compararse con hechos susceptibles de verificación por medio de la observación.

> Según la ley de Hooke, la fuerza necesaria para producir una distorsión en un objeto elástico (tal como un resorte de acero) es directamente proporcional a la cantidad de la distorsión. Se ha observado que un determinado resorte se alarga en una pulgada cuando se le aplica una fuerza de cinco libras. Se aplica una fuerza de diez libras. Síguese que el resorte se alargará en dos pulgadas. Si esto es efectivamente así, el resultado constituye un caso confirmador de la ley de Hooke. […] La hipótesis es una generalización universal; nos permite *deducir* que el resorte experimentará un alargamiento de dos pulgadas. [Salmon 1973, p.119]

O bien, pueden ser generalizaciones estadísticas que son inferidas de forma inductiva[70] como, por ejemplo, en la predicción del clima, conforme se vea que el clima va cambiando a través de los días se registra este cambio para calcular el probable clima que habrá mañana o durante la semana.

Ahora bien, las leyes o generalizaciones universales que son inferidas de forma deductiva tienen predicciones de consecuencias deductivas (hipótesis), éstas son llamadas por Salmon *predicciones de observación*. La clase de argumento que se utiliza para corroborar o desvirtuar (rechazar) una hipótesis se llama: 'método hipotético-deductivo' que consiste en la hipótesis, deducción de las consecuencias de la hipótesis y control mediante la observación para saber si las consecuencias o predicciones de observación son ciertas.

El argumento que lleva de la hipótesis a la conclusión de predicción de observación tiene como premisas: la predicción de observación, las condiciones iniciales y/o hipótesis auxiliares[71]. Sin embargo, cuando esta forma de argumento es aplicada a hipótesis científicas surgen varias dificultades; la primera se origina cuando nos aferramos a nuestra

---

[70] En la ciencia se distinguen clases de argumentos inductivos que ocurren cuando la evidencia se deriva desde la observación, nosotros veremos más adelante dos tipos más conocidos: inducción ampliativa e inducción enumerativa (sumativa).

[71] Las condiciones iniciales son consideraciones que se afirman como verdaderas para apoyar la hipótesis en sus consecuencias, pues especifica circunstancias y factores que deben suceder para que se cumpla la predicción de observación; las hipótesis auxiliares son teorías o leyes que ya fueron corroboradas; son tomadas como verdaderas e incuestionables para apoyar la hipótesis.

hipótesis, pues cuando notamos que no se cumplen las predicciones de observación, preferimos cuestionar si las condiciones iniciales fueron correctas o si hay un error en la predicción observada. Salmon considera que este procedimiento de seguir fieles a nuestra hipótesis por lo general no es ventajoso.

Otro problema se origina cuando tenemos varias hipótesis alternativas que llegan a la conclusión de una misma predicción de observación. Un ejemplo sencillo de entender sobre este problema es el siguiente:

> Se dice a menudo a los niños que las verrugas pueden eliminarse frotándolas con cebolla. Tales "curas" dan a menudo resultado. La hipótesis es que el frotar con cebolla cura las verrugas. Las condiciones iniciales son que Jhonny tiene verrugas y las frota con cebolla. La predicción de observación es que las verrugas de Jhonny desaparecerán. Las verrugas de Jhonny desaparecerán, efectivamente, y esto constituye un caso confirmador de la hipótesis. [Salmon 1973, p.126]

Ahora bien, Salmon se pregunta ¿ha sido corroborada la hipótesis por esta observación? Pero antes de contestar, también hay que considerar que existen explicaciones alternativas de la cura; una de ellas es que se sugiere que las verrugas son síntomas psicosomáticos y que el paciente puede curarse por sugestión, si es así y no se sabe de esta hipótesis alternativa, entonces cualquier hipótesis alternativa puede ser corroborada; por ejemplo, la hipótesis de frotar uvas en vez de cebollas, también puede ser corroborada, pero si la hipótesis explica la cura por síntomas psicosomáticos, es decir, creer que se va a curar, se considera que frotar las verrugas con cualquier cosa ayuda.

En el ejemplo de Salmon, aunque parezca simple, se pueden inferir varios problemas relevantes en la corroboración o rechazo de hipótesis. Si bien, las hipótesis alternativas pueden ser seleccionadas aplicando algún método de la probabilidad, la pregunta sería ¿Cuáles son las probabilidades de que la predicción de observación deducida apoyara una hipótesis *x* y no a otra hipótesis *y* que dan explicaciones alternativas sobre lo observado? Para Salmon éste es el problema más importante de la lógica de la corroboración o rechazo de hipótesis científicas. Ahora bien, el problema aumenta cuando las condiciones iniciales se sustituyen por condiciones iniciales alternativas; si sucede esto, tendremos varias hipótesis alternativas y, por ello, puede haber un número infinito de hipótesis que pueden ser corroboradas por cualquier cosa.

Una solución a este problema es utilizar la probabilidad para elegir la hipótesis que tiene más probabilidad de ser cierta. Para determinar que un caso corroborador apoya una hipótesis es necesario evaluar la probabilidad previa de la hipótesis, esta probabilidad es independiente de sus casos corroboradores y son considerados caracteres plausibles de las hipótesis para calcular la probabilidad de la hipótesis alternativa. De esta forma, el problema es si una hipótesis es inverosímil o absurda, tendrá una probabilidad muy baja, por ello, si tiene sólo un caso corroborador no proporciona un apoyo a la hipótesis, debe de tener muchos casos que corroboren la hipótesis; en cambio para una hipótesis que no es absurda e inverosímil basta con pocos casos corroboradores para que tenga una alta probabilidad.

Sin embargo, el problema de elegir las hipótesis por probabilidad subjetiva es que no se consideran varias hipótesis radicales que bien pueden ser ciertas. Por ejemplo, en el caso de la teoría heliocéntrica contra la teoría geocéntrica; ésta última era más probable subjetivamente, pues iba a favor de las ideas de aquella época y la probabilidad subjetiva está influenciada por los grados de creencia que hay en la persona que hace el cálculo probabilístico, en este caso quien hace el cálculo aritmético puede obtener como resultado un número que puede ser interpretado de manera que favorezca su creencia. En cambio la teoría heliocéntrica el único apoyo que tiene es que debe de tener muchos más casos corroboradores que otra hipótesis (teoría geocéntrica) que no es considerada absurda.

De esta forma, para que sea inductivamente correcto, el método hipotético-deductivo ha de asumir la siguiente forma:

> La hipótesis posee una probabilidad previa no desdeñable.
> Si la hipótesis es cierta, la predicción de observación es cierta.
> La predicción de observación es cierta.
> Ninguna otra hipótesis resulta fuertemente confirmada por la verdad de esta predicción de observación; esto es: las demás hipótesis de las que la misma predicción de observación constituye un caso confirmador tienen probabilidades previas más bajas.
> La hipótesis es cierta. [Salmon 1973, p. 132]

Por ejemplo, en medicina podemos observar un ejemplo del uso del método hipotético deductivo en las pruebas diagnósticas que se realizan y que permiten, conociendo la prevalencia de una herencia genética *y* en una población en la que pertenece un individuo en específico *x* y la especificidad de la prueba, calcular la probabilidad de que este individuo que ha dado positivo en las pruebas, verdaderamente tiene la herencia genética *y*.

Ahora bien, si llamamos *P* a la probabilidad (*a priori*) previa no desdeñable que tiene la afirmación de que la familia *x* tenga la herencia genética *y*, luego, si llamamos su complementaria *Q* y *E* a la especificidad (condiciones iniciales) de la prueba *T*. La hipótesis sería: el individuo *x* tiene la herencia genética *y;* la predicción de observación sería: de las pruebas realizadas todas dan positivo a la herencia genética *y* en el individuo *x*. luego las demás hipótesis que puede haber, en este caso, el individuo *x* no tiene la herencia genética *y*, sino la herencia genética *z*, puede resultar corroborada por la verdad de esta predicción de observación pero tiene una probabilidad previa más baja que la hipótesis que queremos corroborar.

De esta forma, se esquematizan varios argumentos científicos. Las varias hipótesis alternativas, que son muchas, se dice que son plausibles si abarcan un grupo particular de hechos observados. Los problemas que podrían surgir bajo este criterio de eliminación serían; por ejemplo, cuando dos hipótesis distintas tienen la misma probabilidad. Salmon asegura que se puede solucionar el problema de elegir una de ellas, cuando se den algunas situaciones en las que proporcionarán predicciones de observación distintas y en tales casos se practica una prueba de careo para examinar una de dichas situaciones[72] y, de esta forma, saber cuál de las dos hipótesis da la predicción de observación. Sin embargo, podemos objetar que no sabemos cuándo sucederán tales hechos que nos permitirán elegir una de las dos hipótesis[73].

Hay que insistir en que el procedimiento entero de la corroboración de hipótesis científicas es inductivo, es decir, que ninguna hipótesis es verificada como absolutamente cierta; la única garantía de la corroboración de las hipótesis científicas son las varias verificaciones de la hipótesis en una gran variedad de condiciones. Para Salmon ésta es la forma de asegurar la corroboración de la hipótesis. Aunque la verificación de una hipótesis implica varias dificultades, una de ellas y la que nos interesa destacar en este texto es la justificación del método utilizado para verificar una hipótesis.

Justificación de la inducción

En filosofía de la ciencia, la inferencia científica ha tratado con un problema esencial que se relaciona con los métodos científicos que son usados para corroborar o invalidar hipótesis,

---

72 En este párrafo se supone que Salmon considera que hay experimentos cruciales los cuales deciden qué hipótesis debe ser corroborada, sin embargo si alguna de las dos hipótesis tiene más probabilidad, entonces podemos decir que la hipótesis que tiene más probabilidad es la que será corroborada.

73 Éste y otros problemas son objeciones que Gillies atribuye al teorema de Bayes, pero por el momento sólo nos limitaremos a mencionarlos, más adelante se explicarán.

este problema comienza desde que Hume (1748/2004) cuestiona la justificación de la inducción[74]. Antes se creía que los métodos de las ciencias se justificaban por el resultado que se obtenía de ellos, es decir, si en el pasado ha funcionado tal método científico, entonces se cree que usar en un futuro tal método está justificado y se obtendrán mejores resultados que usando otro método que no ha sido usado.

Sin embargo, tales justificaciones al método científico llevaron a varias discrepancias que originaron confusiones. Una confusión fue una mala interpretación de este tipo de argumento en el ámbito deductivo de la lógica, esto llevó a considerar el argumento como una falacia deductiva y consecuentemente tratar de justificar el método con el mismo método que tratamos de justificar. El problema que Hume afirma era no contar con una justificación de la inducción, método utilizado en las ciencias para corroborar hipótesis científicas, pues no podemos justificar la inducción utilizando la inducción, caeríamos en el error de querer justificarlo por medio de ella misma; tampoco podemos utilizar lógica deductiva para su justificación; el problema de la justificación de la inducción es visto como un problema que genera varias perspectivas[75].

El primer error que se cometió desde el principio fue tratar el problema de la justificación del método científico desde la lógica deductiva, pues las situaciones que

[74] Este problema es uno de los más discutidos en la filosofía de la ciencia, en los dos siguientes temas se expondrá el problema.

[75] La primera de ellas es la de Popper, para él la inducción no juega ningún papel en la construcción del saber científico: “[…] se opuso al inductivismo, él afirmaba que el inductivismo es un mito.” [Aliseda y Gillies en Kuipers 2007, p.462] En efecto, Popper afirma que ‘la inducción es un mito’, además afirmaba que la inducción ampliativa nunca podrá ser un procedimiento racional; en cambio, Cohen considera que sí puede ser racional en cuanto, ésta esté conformada de apropiados principios ya sean Baconianos o Pascalianos. La segunda postura, supone que es imposible justificar la inducción y se afirma que nunca ha sido necesaria una justificación; arguyen que no necesita defensa porque es indefendible. La tercera postura, en contraste, sugiere que aunque no existiera una justificación de la inducción, de ello no se sigue que no necesite una justificación.

estudian son diferentes. En el caso del método científico trata de concluir más de lo que tiene en las premisas y por lo general estudia fenómenos falibles y utiliza el método inductivo que trata este tipo de situaciones; además en el método inductivo la conclusión es puesta a prueba para ser corroborada o invalidada y no aparece contenida en las premisas del argumento, a diferencia de lo que sucede en los argumentos deductivos. En el método inductivo la hipótesis es tentativamente corroborada por la observación de la evidencia, si tiene éxito, pero si no tiene éxito, es invalidada.

Esto no significa que se está negando la estructura deductiva que es una parte fundamental en las inferencias en general, pero se exige distinguir las inferencias científicas de las deductivas para evitar confusiones y justificaciones erróneas.

El método hipotético deductivo, el método de inducción por enumeración y el método de inducción ampliativa, son tipos de razonamiento distintos, pero todos implican un procedimiento inductivo. En el caso del método hipotético deductivo se considera que es una inferencia inductiva en el sentido de que hace predicciones y se aplica a hipótesis que no están implícitas en las premisas de forma cierta, además las hipótesis que se concluyen en el método científico son sólo probables.

Cuando se usa el método hipotético-deductivo como método científico se usa para inferir una conclusión inductivamente, pues la conclusión es una hipótesis y no está implícita o explícita como verdadera en las premisas. Lo que se hace con el método hipotético deductivo es usarlo para efectuar inferencias científicas y apoyar una hipótesis dada la evidencia; en las inferencias científicas la hipótesis es el resultado de una investigación científica, por ello, la hipótesis debe ser tratada como una conclusión afirmada que es fundamentada por la evidencia y no como una conclusión que se infiere de las premisas como algo verdadero. La inferencia se hace desde la evidencia observada hacia la hipótesis y esto no es un carácter propio de la deducción, sino de las inferencias científicas que son inductivas; por ello, bajo este uso del método, se dice que un método hipotético deductivo es una inferencia inductiva, además sus hipótesis son sólo probables.

El método hipotético deductivo, según Salmon, que es utilizado en el ámbito científico procede de la siguiente manera: a partir de una hipótesis general y condiciones iniciales que son afirmaciones particulares aceptadas como verdaderas, se infiere una particular afirmación predictiva. Por observación determinamos si la hipótesis afirmada es corroborada o invalidada. Se dice que es corroborada cuando la consecuencia predictiva es verdadera o tiene éxito y si la consecuencia predictiva fuera falsa, la hipótesis sería invalidada. De esta manera, la hipótesis debe ser tratada como una conclusión que está fundada en la evidencia; no debe ser tratada como una premisa más de un argumento deductivo. La inferencia científica, por lo general, se hace desde la evidencia observada hacia la hipótesis, esta forma de proceder no es deductiva.

Por último, la aclaración de las inferencias deductivas e inductivas no responde a la justificación de la inducción, sólo aclara el problema, en la actualidad hay varias formas que tratan de resolver el problema, una de ellas es la aproximación probabilística.

Una justificación probabilística de la inferencia inductiva

La postura probabilística en la investigación científica asume que ningún resultado que se obtiene después de aplicar un método científico es certero, es decir, ninguna hipótesis es establecida con absoluta certeza; el empirismo científico está de acuerdo en que no obtendremos certeza. De esta forma la probabilidad no nos asegura la verdad de una hipótesis, sino sólo nos dice que la hipótesis es probable, y desde una postura subjetiva de la probabilidad, no sólo nos dice que una hipótesis es sólo probable, también nos dice que es menos probable, más probable o ligeramente probable. Por consiguiente, en la interpretación probabilística las inferencias inductivas tienen conclusiones probables; siendo así, el problema que vimos antes sobre la justificación de la verdad de las hipótesis, ahora es formulado como el problema de justificar que una hipótesis sea muy probable o poco probable.

Independientemente de qué teoría de la probabilidad usemos para tratar el problema de Hume sobre la justificación de la inducción, no tenemos razones para suponer que las conclusiones probables sucederán a menudo y que las poco probables sucederán muy poco, pues el argumento de Hume muestra que las inferencias inductivas no siempre conducen a conclusiones verdaderas, por consiguiente no podemos justificar en qué casos no conducen a conclusiones verdaderas, al igual que no sabemos en qué casos serán poco probables: "Cada inferencia inductiva con premisas verdaderas es probable que tenga una conclusión verdadera"[76]. No podemos justificar la afirmación de que cualquier inferencia inductiva con premisas verdaderas tendrá una conclusión verdadera.

En la probabilidad se sabe que una hipótesis se fundamenta en la evidencia; la observación de la evidencia puede decir qué hipótesis puede ser muy probable o poco probable; pero la observación de la evidencia está fundada en la inducción, pues la evidencia que se trata es inductiva y, por lo tanto, la evidencia inductiva se relaciona con las reglas de la inferencia inductiva, por consiguiente, el problema de la inducción ahora

[76] [Hacking 2001, p.250]

puede reformularse como el problema de la evidencia. ¿Qué reglas debemos adoptar para determinar la naturaleza de las pruebas de la evidencia inductiva? Debemos reconocer que el problema es sobre la justificación de la inducción. Cuando tenemos el problema de la evidencia sabemos que los métodos de probabilidad que podemos usar para calcular la probabilidad de una hipótesis dada la evidencia son pocos, el método bayesiano de la probabilidad trata esta observación de la evidencia y los problemas que se infieren de ésta; en cambio, el método frecuencialista de la probabilidad no trata estos problemas, pues parte de eventos que son evidencia para apoyar una hipótesis sin ocuparse de lo que está relacionado con la observación de los eventos, es decir, las posibles discrepancias entre distintas perspectivas que puede haber de los eventos. Más adelante veremos con Hanson este problema de la observación de eventos que son considerados como evidencia.

En consecuencia, veremos cómo la probabilidad sólo nos puede dar reformulaciones del problema de Hume, pero no una solución. Hume muestra que no tenemos razones para suponer que las conclusiones muy probables serán verdaderas y que las poco probables serán muy pocas veces verdaderas.

El problema del concepto de probabilidad

Un problema fundamental que tiene la probabilidad en general es la explicación del concepto de probabilidad, independientemente qué postura adoptemos de la teoría de la probabilidad todas tienen el mismo problema: no hay una noción o definición satisfactoria de lo que debe ser considerado como ‘la probabilidad’.

Como vimos en capítulos anteriores, los axiomas de la teoría axiomática de la probabilidad hacen referencia a la probabilidad de una forma abstracta[77]; los teoremas pueden ser deducidos y los axiomas son supuestos como verdaderos, pero el problema particular de los axiomas de la probabilidad es el significado de sus términos, es decir, la interpretación de la probabilidad. En la actualidad podemos encontrar varias interpretaciones del cálculo de la probabilidad (axiomas de la probabilidad) que deben de justificar las aplicaciones de la probabilidad empírica de la ciencia. Todas las interpretaciones deben dar una explicación rigurosa y formal de la noción de probabilidad que adopten, este problema se relaciona con la estructura lógica de la ciencia y, como dice Salmon, se necesitan conceptos que cumplan diversas funciones dentro de esta estructura; deben ser creados y justificados, el problema es su justificación.

Hay quienes presentan distintos criterios para reconocer una interpretación adecuada de la probabilidad, como Salmon. Entre los varios criterios que podemos encontrar, Salmon presenta tres criterios para que él considere qué interpretación de la probabilidad se adecua a sus criterios para considerarla como una la teoría de la probabilidad, aunque considera difícil encontrar una interpretación de la probabilidad que satisfaga sus tres criterios. El primero es *admisibilidad*, el cual ya explicamos anteriormente (en el capítulo dos); en el criterio de admisibilidad es importante que los conceptos satisfagan las relaciones matemáticas especificadas por el cálculo de la probabilidad (probabilidad axiomática), ya que si no lo hacen tendremos como resultado sistemas incoherentes. El segundo criterio es

---

[77] Salmon en su texto *Logic* considera que una interpretación consiste en una asignación de significados a los términos primitivos. Dos clases de interpretación son posibles una de ellas es llamada abstracta, ésta es la que hace al sistema significativo para la referencia a otras ramas de las matemáticas o de la lógica y esto hace las fórmulas dentro de las afirmaciones entidades abstractas en ese dominio.

aserción o averiguabilidad (ascertainability), en éste se exige que se tenga un método en el cual, por lo menos, podamos afirmar valores de probabilidades. Este criterio expresa el hecho de que un concepto de probabilidad es usado. El tercero es *aplicabilidad*, éste afirma que 'la probabilidad es la guía de la vida', famoso aforismo del obispo Butler. Como una guía de la vida, Salmon, considera que la probabilidad debe ser práctica, es decir, el concepto de probabilidad que busquemos debe tener capacidad predictiva razonable. Por ejemplo: conocemos las probabilidades asociadas a los lanzamientos de los dados que deben tener una importante influencia sobre los tipos de apuestas que los jugadores están dispuestos hacer. Sabemos de la probabilidad asociada a la desintegración radioactiva que debe tener alguna influencia sobre nuestra predicción de la cantidad de una sustancia determinada que permanecerá cierto tiempo sin desintegrarse.

El criterio de aplicabilidad sólo llama la atención a estas aplicaciones prácticas, algunas de las aplicaciones que nuestra explicación de la probabilidad debe satisfacer. Ahora bien, una respuesta por la cual no se puede ajustar a este criterio es simplemente que no hay una explicación del concepto que se trata de explicar.

En conclusión, independientemente de los criterios expuestos por Salmon para obtener una adecuada interpretación de la probabilidad, siempre se deben de satisfacer los axiomas de la probabilidad para construir un concepto de la probabilidad riguroso y formal, es decir, para que sea aceptado como un adecuado concepto de la probabilidad. Los criterios de Salmon son relevantes para poder ver los problemas que surgen en el uso de un método científico.

3.2 Inducción aditiva e inducción ampliativa

Cuando en un razonamiento se infiere más de los datos que se tiene, este tipo de procedimiento, por lo general, se le nombra *inducción ampliativa*, y como vimos anteriormente, es falible.

En contraste, hay otro tipo de inducción que se llama 'aditiva'. Este tipo de inducción establece generalidades sobre los datos que tiene como soporte; el conocimiento que tiene como premisas es el mismo que puede inferir y sólo puede hacer instancias de lo que ya conoce; por citar, el problema de la inducción que refiere Hume con la proposición: 'todos los cuervos son negros', si sabemos que los cuervos que hasta ahora hemos visto son negros, entonces inferimos que todos son negros; pero si uno fuera albino podríamos erróneamente no incluirlo entre lo que consideremos como cuervo; este tipo de inducción también es considerada como aquella que sólo justifica el grado de conocimiento que se tiene.

De este modo, la diferencia entre inducción aditiva e inducción ampliativa es que la inducción aditiva sólo infiere el mismo conocimiento que se tiene implícitamente y la otra inducción infiere más de lo que se tiene como datos (conocimiento que tiene como soporte). La inducción ampliativa es de mayor interés en filosofía de la ciencia, pues fomenta el progreso de la ciencia. Pero, ¿cómo opera la inducción ampliativa? ¿Por qué puede dar más conocimiento que la inducción aditiva? Una respuesta inmediata sería, porque la inducción ampliativa no se limita con tener instancias positivas, es decir, multiplicidad de ejemplificaciones. En el ejemplo anterior de 'todos los cuervos son negros', sólo se llega a la conclusión con la evidencia que nos otorgan los casos que favorecen la hipótesis, no toma en cuenta la evidencia que es negativa o que pueda falsificar la hipótesis; éstas son excluidas, pues se utiliza el método de corroboración.

En conclusión, la inducción aditiva y la inducción ampliativa, ambas tienen el problema que cualquier inducción tendría, a saber, su justificación. La base sobre la cual edificamos nuestras creencias no está justificada; aunque no se pueda demostrar, no significa que no necesite una demostración, podemos plantear problemas acerca de la justificación de cualquier inferencia inductiva. Hemos visto que las teorías de la probabilidad no pueden dar una solución a la justificación de la inducción, pues tiene el problema de su definición de probabilidad, por ello, no podemos decir que los métodos de probabilidad que nos interesa destacar en este texto (método bayesiano y frecuencialista) sean de gran ayuda. Pero como habíamos dicho, en la probabilidad las hipótesis se apoyan en la evidencia para ser corroboradas; luego, la observación que hagamos sobre la evidencia puede afectar en la probabilidad que se le asigne a la hipótesis. Después, la observación de la evidencia está fundada en la inducción, pues la evidencia que se trate es inductiva, por consiguiente, la evidencia inductiva se relaciona con las reglas de inferencia inductiva, por ende, el problema de la inducción ahora puede reformularse como el problema de la evidencia. Este problema de la evidencia, es el que vamos a ver enseguida; también veremos que afecta tanto a los métodos científicos como a los métodos de probabilidad cuando son aplicados y obtenemos resultados de su aplicación; además veremos qué método de la probabilidad los soluciona o por lo menos los trata.

3.3 El problema de la observación de la evidencia y sus efectos en los métodos científicos

3.3.1 El problema de la interpretación de la evidencia

Introducción

Un problema en filosofía de la ciencia es el de cómo se debe interpretar la noción de evidencia. Este problema afecta de forma considerable las razones que formulemos a favor o en contra de un conjunto de hipótesis. Las opiniones o razones pueden afectar de forma considerable la probabilidad de ciertas hipótesis que son estudiadas. Para ser más específicos, el problema trata sobre la observación de los eventos; por ejemplo, un evento *x* puede ser considerado como evidencia para algunos científicos, pero otros no lo verán de ese modo, para ellos podría ser irrelevante. Uno de los problemas de fondo es que puede involucrar la manipulación de los datos para fines personales, ya que los eventos pueden ser interpretados como evidencia a favor de tal hipótesis o pueden ser interpretados como evidencia en contra de la misma hipótesis.

Este problema, para Hanson, (1977)[78], está estrechamente relacionado con la perspectiva, es decir, el punto de vista desde el cual se analiza un evento. Las perspectivas tienen mucha influencia en las controversias y discrepancias que hay entre los científicos sobre cierto evento *x* que observan; por ejemplo, para algunos científicos antes de la teoría heliocéntrica, el sol se movía alrededor de la tierra, pero para otros era la tierra la que se movía alrededor del sol; por ello, un punto a destacar sería saber por qué un evento puede percibirse de manera diferente. Es importante dar cuenta de ello en cualquier observación que se haga, pues puede ser un punto clave para evitar que la evidencia esté influenciada por creencias, o prejuicios; también es importante para decidir sobre temas controvertibles; por ejemplo, el tema del aborto, que puede ser visto desde una perspectiva científica fisiológica y también puede ser visto desde una perspectiva moral. Para Hanson la solución

---

78 "Observación" constituye el capítulo I del libro *Patterns of Discovery And into the Conceptual Foundations of Science*, de N.R. Hanson, publicado por Cambridge University Press, 1958. Versión castellana de García Camarero E. publicada por Alianza Universidad, Alianza Editorial, Madrid, 1977, con el título *Patrones de descubrimiento*.

sería dar cuenta de cómo en la acción de ver ya hay implícita una carga teórica: "En cierto sentido, la visión es una acción que lleva una 'carga teórica'. La observación de *x* esta moldeada por un conocimiento previo de *x*."[79] Pero, más bien, en la percepción ya hay implícita una carga teórica.

A qué se refiere la pregunta "¿Ven el mismo objeto?"

Podemos entender la pregunta como los fisiólogos; para ellos, ver el mismo objeto es dar cuenta de que éste es el mismo para todos, es decir, no va a cambiar cuando lo dejemos de ver o cuando otra persona lo vea; por ejemplo, si dos personas ven la luna llena, una no verá una luna llena distinta que el otro, sino que será la misma y la luz que proyecte será de la misma intensidad; los hombres verán la misma luna, si tienen una visión normal. Así pues, para los fisiólogos, los factores físicos son condiciones suficientes y necesarias para justificar que ven el mismo fenómeno; por ello, los fisiólogos apelando a los factores físicos afirmaban que los hombres ven el mismo fenómeno y por ende tienen las mismas observaciones. En resumen, si se parte de los mismos datos, se tiene la misma observación.

Sin embargo, limitarnos a los factores físicos para justificar la afirmación de que se ve el mismo objeto no es satisfactorio, porque del hecho de que sean afectados por los mismos datos, no significa de forma necesaria que vean lo mismo; entendemos la pregunta ¿Ven lo mismo? no como una pregunta de *facto*, sino que se refiere a la experiencia que afecta de forma decisiva en la visión, como si los dos hombres del ejemplo anterior tuvieran la misma organización de los datos que les son dados o por los que son afectados. Para Hanson que los hombres partan de los mismos datos no significa que vean lo mismo, porque dos personas pueden estar viendo el mismo fenómeno y tener la misma visión

79 [Hanson 2005, p. 238.]

normal, pero ¿a qué se refieren los fisiólogos con tener una visión normal? Si dieran una respuesta sería ambigua, pues no hay forma de encontrar una definición única sobre una visión normal, ya que la experiencia no lo permitiría, porque algunos considerarían que es una buena definición y otros no.

Si para Hanson ver lo mismo significa que tienen la misma organización de los datos que les son dados, es decir; por ejemplo, más adelante (en la figura 1) una persona puede ver algunas líneas del dibujo como el pico de un pato, otra persona puede ver las mismas líneas pero como las orejas de un conejo. Pero, si es posible que veamos cosas distintas al contemplar un mismo objeto o fenómeno, también puede ser a causa de la experiencia y la carga teórica que tenemos influyen mucho en nuestra visión. Ahora bien, si suponemos que los datos son los mismos para todos, pero la organización de ellos es distinta, ¿cómo son los datos moldeados u organizados por diferentes interpretaciones que llevan a construcciones de evidencia a favor o en contra de cierta hipótesis científica *x*?

La cuestión de fondo es importante porque afecta a las decisiones de los científicos, de aquí las preguntas que se siguen son para Hanson: ¿Ven lo mismo? ¿Cómo es posible ver diferentes cosas viendo un mismo objeto? Podemos plantear la pregunta desde otro punto: ¿Qué es lo que ambos perciben en un objeto?

Creo que para entender la postura de Hanson, es necesario saber que su pregunta no es ¿Ven el mismo objeto físico? más bien, su pregunta es: ¿Cómo es posible que las personas tengan distintas interpretaciones de un mismo objeto?

Por ejemplo, dos personas discrepan al percibir un mismo fenómeno natural, por un lado, uno de ellos afirma que la tierra se mueve y la luna está fija; en cambio, la otra persona asegura que la luna se mueve y la tierra está fija ¿Acaso los dos hombres perciben el mismo evento? Como ya habíamos dicho, para Hanson la pregunta ¿Ven lo mismo?, es una cuestión de observación que supone la experiencia, pues ésta influye en la forma en que vemos las cosas. Sin embargo, en los siglos XVI y XVII, no era considerada así. La cuestión de Hanson no es de *facto,* sino más bien, el comienzo de un examen de los conceptos de visión y observación. La discusión que genere la respuesta a la pregunta ¿Ven lo mismo? puede dar para una extensa discusión.

Ahora bien, si ya habíamos dicho que dos hombres pueden ver lo mismo, por las razones de que son afectados de la misma manera y tienen la misma visión normal, esto es, lo que afirmaban los fisiólogos, pero Hanson replica que:

> La visión es una experiencia. Una reacción de la retina es solamente un estado físico, una excitación fotoquímica. Los fisiólogos no siempre han apreciado las diferencias existentes entre las experiencias y los estados físicos. Son las personas las que ven, no sus ojos. [Hanson 2005, p. 228]

La visión es una experiencia que no depende sólo del estado físico, si sólo fuese eso, no habría diferencias y no tendría caso esta reflexión; son las personas las que ven no sus ojos, los ojos sólo posibilitan fisiológicamente la visión, por ello, decir que dos hombres están viendo la misma cosa, no puede argumentarse mediante las referencias a estados físicos de sus retinas; lo que si hay que preguntarse es dadas las condiciones de las personas cómo es posible que tengan distintas interpretaciones de lo que perciben.

Para Hanson decir que los dos hombres tienen en cierta forma una experiencia visual común, es decir que esta experiencia quizás constituye su ver la misma cosa; por ejemplo, dos hombres pueden estar de acuerdo en que el sol es un círculo brillante, pero puede suceder que dos hombres cuando dicen que ese es el sol no estén significando lo mismo. Esto es porque, cuando dicen “veo el sol” quizás este enunciado tiene distintos significados para ellos que enuncian la misma oración, ésta es, una de las formas de ver el problema: ver distintas cosas en un mismo objeto, es decir, significar distintas cosas con el mismo enunciado; de ello podemos inferir que no están viendo la misma cosa, en el sentido de decir que la visión es una experiencia; aunque estuvieran enunciando y contemplando el mismo objeto (un estado físico) no podemos asegurar que están viendo lo mismo. Por ello, no podemos decir que es un buen criterio para justificar que están viendo lo mismo, cuando dicen lo mismo ante el objeto o fenómeno que contemplan.

Con respecto a la pregunta: ¿están viendo lo mismo? Si dos personas ven un mismo dibujo, como la figura 1 que es considerada como un *gestalt-switch*, tal vez no se pueda describir la diferencia en que vean distintos objetos y por ello, también, que no puedan describir que ven el mismo objeto.

> Pero también, de nuevo, podemos decir que no vemos la misma cosa (El sentido en que podemos ver la misma cosa empieza a perder su interés filosófico). Yo dibujo una copa. Usted dice: “esto es justamente lo que yo veo, dos hombres que se miran fijamente”. ¿Qué pasos hacen falta para conseguir que usted vea lo que yo veo? ¿Es que la imagen visual de uno cambia cuando la atención se desvía desde la copa hacia las caras? ¿Cómo cambia? ¿Qué es lo que cambia? ¿Qué puede cambiar? Nada óptico o sensorial ha modificado, y, sin embargo uno ve cosas diferentes. Cambia la organización de lo que uno ve. [Hanson 2005, p.228]

¿Cómo justificar y describir de una forma necesaria y suficiente que uno ve lo mismo que el otro? ¿En qué estamos coincidiendo? ¿Cómo saber que estamos coincidiendo en algo?

Para Hanson, decir que tienen una experiencia común es decir que están coincidiendo en algo que involucra algo de lo que no están coincidiendo, es decir, que los dos hombres, del ejemplo anterior, pueden estar de acuerdo en que el sol es un círculo brillante y que es un cuerpo celeste, pero no están de acuerdo en que se mueve o no; ¿perciben la misma cosa? Podríamos decir que no y dar varias razones, pero nuestra respuesta sería muy ambigua. Ahora bien, si como dice Hanson, cambiamos nuestra perspectiva y preguntamos ¿qué es lo que ambos ven? ¿Podríamos obtener una respuesta menos ambigua?

Si bien, los dos hombres del ejemplo tienen fija su atención en un disco brillante (el sol). Tal imagen de "datos sensoriales" es única y no invertida, es decir, están de acuerdo en algo, en que el sol es un disco brillante y lo que se discute es si se mueve o no se mueve (con relación a la Tierra). Su discusión está relacionada con estar de acuerdo en algo del fenómeno observado. Así, lo único que nos indica que estamos viendo el mismo fenómeno es estar de acuerdo, en este caso, que es un cuerpo celeste brillante, pues al indicarnos en qué estamos de acuerdo y no hay objeción en ello, también nos indica al mismo tiempo en qué estamos en desacuerdo. Pero ocurre un problema cuando se enuncia 'está de acuerdo con *x* fenómeno', pues no significa necesariamente que veamos el mismo fenómeno. Se puede objetar esta corroboración y decir que del hecho de que enunciemos "el sol es un cuerpo brillante", no garantiza que estamos de acuerdo en que vemos el mismo fenómeno, si es así, ¿qué nos garantiza que vemos el mismo fenómeno?

Hasta aquí, Hanson nos ha indicado que los factores físicos no son suficientes razones para decir que se está viendo el mismo objeto y que el enunciar la misma oración no significa que veamos lo mismo, en cambio, considera que la experiencia y la carga teórica son factores esenciales que determinan lo que vemos y da por hecho que los datos que nos son dados son los mismos (el mismo objeto), pero los organizamos de diferente forma que nos permite ver al objeto de diferente forma; esto es, que nuestra visión está determinada por nuestra experiencia y carga teórica, independientemente de que no encontremos un criterio suficiente y necesario para decir que dos personas están viendo el mismo objeto. Por ejemplo la figura 1 puede ser vista como un pato o un conejo, la persona que sólo ha visto conejos y nunca ha visto un pato, ni algún animal que se le parezca, verá al conejo, en cambio, una persona que ha visto tanto patos como conejos puede ver tanto al pato como al conejo en el dibujo (esto en psicología es conocido como una ilusión óptica).

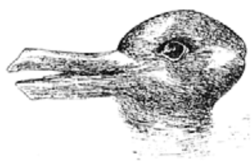


Figura 1.

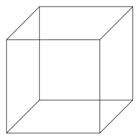


Figura 2.

Ahora bien, suponemos que ven la misma cosa porque tienen la misma experiencia de datos sensoriales. Las disparidades o discrepancias aparecerán en las interpretaciones de lo que se ve, no en los datos visuales básicos. Si se sostiene esto aparecen dificultades como las siguientes. En este ejemplo y hasta ahora Hanson nos ha expuesto la problemática que se relaciona con la observación de la evidencia basándose en la dificultad de justificar que se está viendo el mismo evento; sin embargo su exposición sólo va dirigida a la observación en el sentido de ver un evento, pero para nuestro interés, la visión de un evento es sólo un ejemplo del problema de la percepción de un evento, que no sólo involucra la visión sino más. Sin embargo no podemos detenernos en cada uno de las cosas que involucra la percepción, por ello, sólo expondremos el problema de la observación de un evento. Después de ver a Hanson veremos cómo con Russell el problema de la observación de la evidencia involucra la percepción de los eventos y los datos; pero por ahora continuemos con Hanson.

Interpretación

Una dificultad que puede aparecer es la siguiente: en la figura 2 se puede ver (percibir) un cubo o líneas o un cubo que es visto desde arriba o un cubo que es visto desde abajo, etc. El dibujo puede ser interpretado de muchas maneras, pero se considera que el estado físico es el mismo, es decir, las reacciones de la retina ante el evento (dibujo) es la misma; los datos son los mismos, pero la organización de ellos es distinta, esto se puede entender como la interpretación de los datos. Ahora bien, la pregunta siguiente sería, ¿tener una interpretación diferente del objeto significa que estamos viendo (percibiendo) cosas diferentes? Podemos estar de acuerdo en que es una caja, al igual que podemos estar en desacuerdo en que es una caja vista desde arriba o una caja que se ve desde abajo. ¿Se puede estar interpretando cuándo se está viendo, en este caso, un mismo objeto (una caja)? Para Hanson algunas veces sí es posible, como cuando vemos la cara de alguien podríamos interpretar qué estado de ánimo tiene. Pero, ¿es esta interpretación la que opera cuando se ven claramente las cajas? ¿Opera esta interpretación cuando la caja de repente cambia de caja que es vista desde arriba a caja que es vista desde abajo?

La experiencia común que menciona Hanson supone que hay un conocimiento común que nos indica que vemos lo mismo, aunque no hay justificación satisfactoria de que tenemos una experiencia común: "'Pero, la interpretación se hace en muy poco tiempo, es instantánea'. La interpretación instantánea proviene del Limbo que produce *sensibilia* no sentidas, inferencias inconscientes, enunciados incorregibles, hechos negativos y *Objektive*.” [80]

Hanson considera que la interpretación es instantánea, esto nos lleva a pensar que los datos que nos son dados pueden no ser los mismos, porque cuando vemos un objeto puede existir una experiencia visual que no es común, la interpretación se encuentra al recibir los datos, no esperamos que primero nos den los datos, sino que nosotros inmediatamente interpretamos que esos son los datos. De esta afirmación se sigue la pregunta: ¿cómo probar que estamos partiendo de los mismos datos para dar una interpretación o si ya estamos interpretando los datos?

> En ningún sentido ordinario de la palabra ‘interpretar’ yo interpreto la figura 1 de modo diferente cuando su perspectiva se invierte para mí. No está claro que en el lenguaje ordinario ni en el lenguaje extraordinario (filosófico) exista un sentido extraordinario de la palabra. El hecho de insistir en que las diferentes reacciones ante la figura1 deben descansar sobre interpretaciones hechas a partir de una experiencia visual común es justamente reiterar (sin razones) que la visión de *x* debe ser la misma para todos los observadores que están mirando a *x* […] Las teorías y las interpretaciones están allí, en la visión, desde el principio. [Hanson 2005, p. 226]

Ahora, consideremos que nuestra forma de ver la figura 1 también depende del contexto; por ejemplo: si una persona conoce tanto a los patos como los conejos; pero, si yo hablo de aves, muestro varias imágenes de aves y el tema que es expuesto es sobre aves, es más probable que la persona vea primero en la figura 1 un pato que un conejo, aunque no es necesario que el contexto sea explicitado; la persona debe de entender de que estoy hablando de aves, que estoy hablando de una experiencia que tenemos en común.

Hasta ahora se ha dado una descripción de la visión basada en una concepción de los datos sensoriales, pero hay otras formas de describirlas, una de ellas es apelando al

---

[80] [Hanson 2005, p. 225]

contexto, el contexto es parte de la ilustración porque, para Hanson, ver un pato o un conejo en la figura 1 depende mucho del entorno lingüístico del cual depende el sentido y el valor que se dote a una imagen, pero no todo depende de ello. En cierto sentido, la percepción o como diría Hanson la visión lleva una carga teórica, por eso podemos decir que todos los científicos no hacen las mismas observaciones, pues, si fuese así no habría descubrimientos ni paradigmas al igual que la famosa ruptura revolucionaria que Kuhn nos describe en las revoluciones científicas.[81](Lo que dice Kuhn es que, en efecto, hay una ruptura, aunque, visto desde otra perspectiva, es un seguimiento no una ruptura, es el mismo problema.)

Kuhn explica este cambio de perspectiva por medio de "flashes" de intuición, afirma que cuando cambiamos de perspectiva, el cambio es de inmediato, de eso se tratan las revoluciones científicas: el cambio de perspectiva. La problemática a la que se refiere Kuhn es que cuando aparece un nuevo paradigma no se puede ver a una teoría científica como una corrección de la teoría que le antecede, pues los conceptos involucrados, aunque lleven el mismo nombre, no son iguales, han cambiado significativamente, son *inconmensurables*.[82].

Las controversias existentes en los procesos de búsqueda que hay en el ámbito científico no se pueden explicar por la simple epistemología que afirma que todos los observadores ven la misma cosa *x* pero la interpretan de forma distinta; como dice Hanson: "es importante darse cuenta de que destacar diferencias en datos, elementos de juicio y

---

81 Para Kuhn no es legítima la comparación del cambio de perspectiva, pues tienen distintas concepciones y tratar de comparar algo es imprimirle la postura que asumimos, de esta forma no tomaríamos el paradigma anterior tal y como es, pues de inmediatamente lo interpretaríamos. [Aliseda y Gillies 2007, p.235]

82 Kuhn reconoce que el termino fue tomado de las matemáticas; en aplicación al termino 'inconmensurable' a las teorías científicas, él ha intentado decir, al utilizar el término, que no hay un lenguaje en común dentro del cual dos paradigmas se expresen totalmente y en el cual puedan, por lo tanto ser comparados punto por punto. [Aliseda y Gillies 2007, p.235-238]

observaciones, puede requerir algo más que el simple gesticular ante los objetos observables."[83]

La simple observación de un objeto presupone un conocimiento que afecta la forma en que lo vemos, por eso podemos decir que no hay observaciones 'puras'; nuestras observaciones siempre estarán precedidas por una carga teórica, como dice Moulines: "Los datos se detectan mediante la observación. La observación está vinculada casi siempre a la realización de un experimento, […] pero a veces se observa sin experimentar en sentido estricto."[84] La observación, ya sea experimental o no, tiene presupuesta la carga teórica que Hanson explica en su texto "Observación".

En conclusión, para Hanson el conocimiento tiene su origen en la percepción y en la experiencia de los la persona que adquiere el conocimiento (no necesariamente sólo el lenguaje natural o la visión como Hanson destaca, en el siguiente tema veremos de una forma más general la percepción). No hay interpretación mejor que otra, todas se consideran genuinas, pues detrás de cada percepción hay un conocimiento que crea coherencia en lo que vemos, pero también nos permite percibir las cosas desde otra perspectiva. Es importante resaltar que ninguna forma de percibir las cosas es única, pues no podemos sostener que mi percepción de las cosas es la mejor, porque creo que yo percibo las cosas tal y como son, por ello, quien tenga una distinta percepción de la mía está equivocado. Además debemos resaltar que no hay un objeto o cosa neutra de la que podamos decir que todas las perspectivas son interpretaciones de ella**.**

---

[83] [Aliseda y Gillies 2007, p.238]

[84] [Diez y Moulines 1999, p.74]

Esta distinta forma de percibir es lo que impulsa las revoluciones del pensamiento humano; la observación es importante en las revoluciones científicas, pues una gran parte de la decisión que se tome para saber cómo actuar ante un fenómeno depende de la evidencia y de cómo son vistos los datos. Los datos que se interpretan mediante la observación son importantes para corroborar o refutar las hipótesis que se tienen ante un fenómeno. Por eso la observación de ellos es importante y reconocer que existe un problema cuando no se sabe con exactitud qué debe ser considerado como evidencia a favor o en contra de una hipótesis *x* es sólo uno de los problemas que se discuten en filosofía de la ciencia.

Hemos de mencionar que estas distintas formas de interpretar los datos se deben a nuestro conocimiento del mundo externo. En la siguiente sección veremos cómo Russell hace un análisis al respecto.

3.3.2. Conocimiento basado en la probabilidad

Introducción

El conocimiento que obtenemos del mundo externo nos lleva a preguntar qué es lo que conocemos y cómo es que lo conocemos, éstas son dos preguntas fundamentales qué trata Russell en su texto *Our Knowledge of the external world* (1926). Hay dos posturas, expuestas en el texto de Russell, que tratan de dar una justificación al problema del conocimiento del mundo externo. Por un lado, se encuentra el mundo de la física, así son las cosas independientemente de que las percibamos; por otro lado, el mundo de los sentidos, así son las cosas y dependen de cómo las percibimos. En este texto sólo se discutirá la teoría del mundo de los sentidos. Russell intentará dar razones para creer que la teoría del mundo de los sentidos es débil.

El mundo momentáneo de las sensaciones

Si suponemos que todo lo que conocemos es lo que sentimos, lo que percibimos, entonces nuestro conocimiento se fundamenta en nuestra percepción. Pero, para Russell antes de comenzar por preguntarnos por la percepción debemos preguntarnos por el conocimiento que tenemos, para Russell éste se fundamenta en la evidencia que ésta, a su vez, no es más que la representación de los datos, es decir, en última instancia debemos preguntarnos qué es lo que llamamos dato. Es aquí por donde Russell (1926) comienza. Él considera un dato como algo vago, complejo e inexacto que es materia del conocimiento común. Si los datos que obtenemos a partir de la experiencia particular de los objetos son materia del conocimiento común, ¿cómo es posible que desde experiencias particulares se pueda generar conocimiento común?

Si comenzamos a discutir sobre lo que percibimos, debemos exponer las circunstancias y los factores que influyen en nuestra percepción. Por ejemplo, cuando percibimos una mesa puede ser percibida de distinta forma, pues son posibles los distintos puntos de vista: "Una mesa vista a partir de un lugar presenta una diferente apariencia de la que presenta en otro lugar. Este es el lenguaje del sentido común, pero este lenguaje ya asume que hay una mesa real desde la cual vemos las apariencias." [85] Además de influir la posición en que es vista la mesa sobre nuestra percepción, también influye la fisiología de las personas. Por ejemplo, si una persona que tiene una enfermedad en los ojos ve la mesa, su percepción será afectada de manera distinta a la de una persona que no tiene tal enfermedad. Nuestra fisiología, como las circunstancias en que percibimos la mesa, afectan

[85] Este es el texto original de mi traducción : A table viewed from one place presents a different appearance from that which it presents from another place. This is the language of common sense, but this language already assumes that there is a real table of which we see the appearances. [Russell 1926, p. 84]

nuestra percepción y por ende nuestro conocimiento. Pero, como dice Russell, hay algo común en todas las percepciones: asumir que la mesa es real. El conocimiento común[86] nos permite estar de acuerdo que vemos una mesa, aunque la percibamos de distinta forma.

De todas las percepciones generadas al ver la mesa, sólo podemos reconocer algunas, pues, según Russell, algunas percepciones serán parecidas y coincidirán en que dos personas ven una mesa; habrá otras que no será el caso, pues puede haber, como dice Russell, infinidad de perspectivas de las cuales algunas no las podremos percibir si no son parecidas a nuestras perspectivas. Russell se refiere a la posición espacio temporal en la que estamos y que nos permite observar desde cierto ángulo un objeto y de esta forma obtener una perspectiva única: "El sistema consiste en todas las visiones del universo, las que son percibidas y las que no son percibidas, debemos llamarlo el sistema de las 'perspectivas'."[87]

Ahora bien, Russell menciona que si se quiere justificar nuestro conocimiento de la existencia presente de los objetos que nos dan los sentidos, se tiene que exponer el supuesto de que los objetos persisten a pesar de que dejamos de percibirlos. Si consideramos a los objetos como reales, debe persistir su existencia cuando los dejemos de percibir o no lo hayamos percibido. Por ejemplo, la mesa que tocamos o vemos no deja de existir cuando dejamos de percibirla. "Estos objetos persisten aún en la ausencia de cualquier percepción."[88]

Los objetos sensibles persisten, es decir, los objetos que suceden ahora sólo pueden ser explicados por lo que sucedió anteriormente, de ahí la importancia que Russell le da a la

86 El conocimiento común al que Russell se refiere es más cercano a la noción de sentido común de Thomas Reid (1764) que a la noción de "conocimiento común" de Lewis (1962).

87 [Russell 1926, p.95]

88 [Russell 1926, p.82]

memoria, pues justifica nuestro conocimiento de que los objetos siguen ahí, argumentando que ya estaban antes de percibirlos. Pero Russell se pregunta por lo que sucede con los objetos o hechos que no se pueden percibir; por ejemplo, hechos de la historia. Sabemos que Napoleón existió, sin embargo nunca lo percibimos; tenemos conocimiento de ello por el testimonio de otras personas. Dice Russell:

> Pero que depende del testimonio, sin ser oído o leído, no puede ser explicado desde esta postura, puesto que el testimonio depende sobre la existencia de otros pensamientos más que el de nosotros mismos, y de esta manera requerimos un conocimiento de algo que no es dado en los sentidos. [Russell 1926, p.56]

Si bien conocemos hechos u objetos que nunca hemos percibido o bien percibimos de distinta manera esto no explica cómo es posible que tengamos un conocimiento común. Por ello, Russell expone otros factores que pueden influir en nuestra percepción, por ejemplo, nuestra experiencia afecta a nuestra percepción: si fui educado de cierta forma, voy a ver las cosas de cierta forma. Tal vez aquí tenga que ver el sentido común que no es más que una construcción de un conocimiento que comparte una comunidad, como dice Russell: “Si usted no está acostumbrado al gramófono, con seguridad creerá que la voz que oye al otro lado de la puerta es de una persona […] al mismo estimulo externo, pueden reaccionar de distinta forma dos personas.”[89]

Alguien puede reconocer el sonido de un gramófono, si está acostumbrado a asociar de cierta manera las cosas, es decir, sabe qué es un gramófono y los sonidos que produce; en cambio para otra persona, que no sabe qué es un gramófono, puede pensar que es una persona la que habla. Estamos acostumbrados a asociar las cosas de cierta manera, pues nos

---

[89] [Russell 1926, p.185]

hemos educado en una sociedad o grupo en que nos enseñaron a comprender las cosas de cierta manera.

La costumbre es un factor importante que nos permite explicar por qué nuestro conocimiento de los objetos puede ser común, al igual que explica por qué no es necesario que sean percibidos por nosotros, es decir, no necesitamos una experiencia particular para legitimar la existencia del objeto.

Ahora bien, como dijo Russell, sólo hemos especulado sobre el mundo momentáneo de los sentidos y no hemos construido algún mundo más estable, debemos buscar otros factores que se relacionen con la percepción, factores que justifiquen el conocimiento común que tenemos de los objetos.

De costumbres a sentido común

Según Russell, lo que consideramos como percepción de los objetos externos que existen, consiste en la costumbre generada por experiencias pasadas. Cuando vemos una mesa o un gato, no suponemos que dejan de existir cuando no son percibidos, no tienen una existencia momentánea que depende de nuestra percepción, sino que sabemos que los objetos tienen un pasado y un futuro.

Otro punto que aborda Russell en su texto es sobre el problema de justificar la existencia los objetos. Si bien, al principio ha supuesto que pueden ser justificados porque los percibimos, ¿cómo podemos justificar los objetos que parecen ilusiones en los sueños? Russell dice que por lo general estamos acostumbrados a asociar las cosas de cierta manera, pero lo importante es que encontremos una asociación de las cosas que sea correcta con la realidad, con nuestra experiencia, pues de esta forma se explica por qué en los sueños las cosas que son reales parecen ilusiones; los objetos de los sentidos no son considerados como ilusiones cuando están relacionados con otros objetos. "Los objetos de los sentidos son llamados 'reales' cuando tienen una clase de conexión con otros objetos de los sentidos los cuales la experiencia nos ha llevado a referirnos a ellos como normales; cuando fallan en esto, son llamados 'ilusiones'." [90]

De esta forma, Russell explica cómo los objetos de los sentidos cuando ocurren en los sueños, son los más indudables objetos reales que conocemos; lo que cambia, no son los objetos de los sentidos, sino la relación, la forma en que están relacionados, si es una relación normal[91] no son llamados ilusiones, pero si la relación es inusual con otros objetos de los sentidos, entonces sólo son ilusiones. Lo irreal son las relaciones.

La memoria es importante para fundamentar nuestro conocimiento, pues hace más fuerte nuestras creencias. Por ejemplo, si nosotros recordamos que un hecho sucede de tal forma en ciertos momentos, podemos decir que nosotros sabemos que va a ocurrir tal hecho en tal momento en un futuro, es decir, sabemos qué tipo de relación esperamos que suceda, por eso Russell enfatiza mucho la memoria como un factor importante que justifica nuestro conocimiento del mundo exterior.

---

90 Este es el texto original de mi traducción: Objects of sense are called 'real' when they have the kind of conecction with other objects of sense which experience has led us to regard as normal; when they fail in this, they are called 'illusions'. [Russell 1926, p.93]

91 Entendiendo por normal, aquella relación a la que estamos acostumbrados a ver en los objetos a los que nos referimos y asociamos de cierta manera.

La conclusión a la que llega es que el conocimiento es una materia de grados de credibilidad, pues el conocimiento que obtenemos es para algunos indudable pero para otros no, siempre dudaremos de nuestro conocimiento. Por ejemplo, al percibir los hechos u objetos constantemente estamos dudando de ellos, aunque con un cierto grado de duda. El alto grado de credibilidad es fundado en hechos de percepción y la memoria. El principio general de creencias que se utiliza es, en este caso, el inductivo, pues nos estamos refiriendo al mundo falible, el de la vida diaria.

Nuestro conocimiento es sólo probable

Como ya decíamos al final del segundo capítulo, Russell en su reflexión y crítica de nuestro conocimiento sobre el mundo exterior, llega a considerar la probabilidad como una característica de nuestro conocimiento. Además de que no podemos obtener certeza de datos dudosos, si queremos hacer una inferencia desde premisas que suponemos como verdaderas y con un razonamiento correcto, sólo obtendremos conclusiones probables.

El conocimiento que obtenemos siempre es dudoso, y la probabilidad es la que nos puede aclarar los grados de credibilidad que tenemos sobre nuestros conocimientos. En la vida diaria tenemos conocimientos y creemos en ellos, pero necesitamos aplicarlos para resolver problemas de un mundo falible, por ello la probabilidad nos sirve para que nos aclare y exponga qué conocimiento es mejor aplicar. Por ello, la lógica se ha ampliado para tratar problemas que en la lógica deductiva no eran tratables, pues, eran falibles[92]

Un atributo que Russell resalta de la inducción es el ejemplo del razonamiento de Hume, donde concluye que todos los cuervos son negros de premisas que apelan a la

[92] 'Falible' es usado aquí como una referencia a resultados que se obtienen al aplicar un método y se pueden obtener distintos resultados, pues no se toman en cuentan factores que cambian de forma, a veces radical, los resultados.

experiencia particular; el argumento implícitamente expresa que no se necesita experimentar, percibir, todos los cuervos para llegar a la conclusión de que todos los cuervos son negros; es más, nos es imposible saber de forma perceptiva que todos los cuervos son negros, de ello podemos concluir que todo nuestro conocimiento no se deriva de los sentidos. En este ejemplo, muestra que no todo nuestro conocimiento depende de la experiencia, sino que podemos obtener conocimiento con inferencias. Dice Russell: “Por lo tanto, debemos admitir que el conocimiento general no se deriva desde los sentidos.”[93]

Sin embargo, la inducción tiene varios defectos que Russell expone; por ejemplo, dice que si el proceso de inducción es por enumeración, habrá casos en que no se pueda aplicar, casos que no se puedan enumerar. De este proceso, Russell dice: “consiste en la atribución de la naturaleza de las verdades generales a todas las proposiciones las cuales son verdaderas en cualquier ejemplo o instancia que conocemos”[94]. En lo que concierne en su falibilidad él afirma que: “La poca estabilidad del método de simple enumeración es en una consciente inversión de la generalización. El proceso es engañoso e insuficiente, exactamente en proposiciones como los temas de la observación que hemos estado tratando.”[95]

Hemos de distinguir, como bien señala Russell, entre la probabilidad matemática, cuya aplicación sólo se restringe a casos numerables, y clases de probabilidad que no son matemáticas y que tratan con casos no numerables. Cuando se usa la probabilidad es porque no estamos seguros de nuestro conocimiento y en la vida diaria constantemente utilizamos la probabilidad, por ello, necesitamos de otra teoría de la probabilidad que

---

93 [Russell 1926, p.66]
94 [Russell 1926, p.189]
95 [Russell 1926, p.189]

abarque aquellos casos que no son numerables. La probabilidad, en contraste con la inducción, puede aplicarse a casos que no son numerables. Por lo que hasta ahora hemos dicho de la postura de Russell, podemos decir que él considera que la probabilidad es más amplia que la inducción en su aplicación a casos cualitativos.

Ahora bien, si nuestro conocimiento es dudoso, tenemos ciertos grados de duda sobre distintos conocimientos, es decir, tenemos cierto grado de credibilidad, 'los grados de credibilidad' es otro tipo de probabilidad que abarca aquellos casos que no son enumerativos.

El conocimiento en particular es considerado como el conocimiento que obtenemos de nuestra experiencia personal; se encuentra inmerso en el conocimiento común. El conocimiento común es el que obtenemos de la vida diaria. Sin embargo, al dudar de él no estamos diciendo que sea verdadero o falso, sino, como dice Russell, tiene grados de credibilidad y por ello, creemos en los conocimientos que tenemos de la vida diaria con cierto grado de credibilidad, claro que dependen varios factores que pueden apoyar y hacer más fuerte nuestra creencia o pueden hacerla más débil, es decir, en términos de Russell, menos dudosa o con más credibilidad.

Conclusión

Nuestro conocimiento del mundo externo envuelve preguntas filosóficas que Russell trata apuntando a problemas como el suponer que nuestro conocimiento sólo depende de nuestra percepción. Al final de su investigación regresa a la definición de conocimiento y sugiere varias formas de justificar lo que debemos definir por conocimiento, llega a la conclusión de que preguntarse por eso y preguntarse "¿qué debemos definir por calvo? " tienen la misma dificultad de encontrar una respuesta.

Al final, llega a la conclusión de que se tienen grados de credibilidad siempre que se hable de conocimiento; se tiene que hablar de los grados de credibilidad y clases de probabilidad no enumerativa que Russell expone en su texto.

Ahora bien, a lo largo de todo el capítulo hemos visto, como Russell expuso que sólo podemos aspirar a que una gran parte de nuestro conocimiento y en particular nuestro conocimiento del mundo empírico sólo sea probable. En filosofía de la ciencia sólo pueden ser corroboradas las afirmaciones por medio de la evidencia; si bien la interpretación de la probabilidad afecta en su aplicación para la afirmación de una proposición, lo que Russell ha venido enfatizando es que la probabilidad en la práctica tiene una fuerte relación con los grados de credibilidad y que la probabilidad no se limita a la asignación de números a las proposiciones, sino que hay factores que afectan en su asignación de números a las proposiciones y una de ellas es la evidencia, preguntas como qué debe ser considerado como evidencia, cómo deber ser interpretada y el problema de la inducción de Hume entre otras preguntas que se ven como problemas filosóficos en la filosofía de la ciencia. Ahora bien, en el siguiente capítulo veremos con más detalle la aplicación de los métodos bayesiano y frecuencialista en el método hipotético deductivo científico.

IV. Estimación de la aplicación de los métodos bayesiano y frecuencialista al método hipotético deductivo científico

Introducción

En este capítulo se expondrán sólo algunos métodos científicos que utilizan el método probabilístico bayesiano y el método probabilístico frecuencialista para la corroboración o refutación de hipótesis científicas; se expondrán las ventajas y desventajas de su uso, se darán ejemplos de la aplicación de los métodos de probabilidad, tanto del bayesiano como del frecuencialista, esto último se hará para exponer qué método de la probabilidad tiene una mejor utilidad cuando se usa en algún campo de estudio en específico. Después de saber por qué cierto método es más útil que otro en cierto campo de estudio, se retomarán estas razones y las otras conclusiones que hemos llegado a lo largo de cada capítulo, para poder exponerlas en el capítulo cinco y contrastarlas con el fin de llegar a sustentar la hipótesis principal de todo este texto, a saber, el método bayesiano de la probabilidad es más eficaz en su uso en los métodos científicos para la corroboración de hipótesis que el método frecuencialista de la probabilidad.

4.1. El teorema de Bayes aplicado a la corroboración de hipótesis científicas

4.1.1 El análisis bayesiano en la corroboración de hipótesis científicas

En este capítulo sólo mencionaré los métodos científicos que son más controversiales y usados en la actualidad que usen la regla de Bayes de forma directa para calcular la probabilidad de hipótesis alternativas. Por ello, explicaré el método hipotético deductivo científico que expusimos en el capítulo tres y los argumentos de falsificación y plausibilidad que Salmon usa, para tratar de solucionar un problema que surge cuando se intenta usar el teorema de Bayes para calcular la probabilidad de las hipótesis científicas.

Antes de que se aplicara la probabilidad en la inferencia científica, uno de los problemas esenciales en el método científico era carecer de un criterio formal y riguroso para probar las hipótesis científicas a partir del evento a explicar. Después de la teoría de la probabilidad se usaron varios teoremas de la probabilidad para tener un método formal científico; entre los teoremas que se usan está el teorema de Bayes cuya utilidad consta de calcular la probabilidad de hipótesis alternativas de un fenómeno a explicar; la forma condicional que tiene el teorema de Bayes permite calcular la probabilidad de las posibles causas del fenómeno, además puede adaptarse bien al método hipotético deductivo.

Otro problema que enfrenta el método científico es la variedad de pruebas que se deben aplicar a las hipótesis. Si recordamos el ejemplo de Salmon en el capitulo tres, él concluye que corroborar una hipótesis por medio de sólo un experimento es insuficiente, pues limitarnos a un experimento puede corroborar a varias hipótesis; pero si usamos un método formal que nos permita considerar varios experimentos y de esta forma tener un mejor criterio para probar las hipótesis, se podría considerar como un método eficaz.

Por un lado, el método bayesiano nos permite realizar varios experimentos y eliminar de una manera rigurosa y formal más hipótesis; por otro lado, el método hipotético deductivo científico puede usar el método bayesiano para calcular la probabilidad de las hipótesis y, de este modo poder restringir su atención a las hipótesis más probables para después aplicar experimentos e ir descartando las que son menos probables y obtener más evidencia a favor de una de las hipótesis restantes.

El método de Bayes calcula la probabilidad de una hipótesis de una forma causal, es decir, la hipótesis que trata de probar es considerada como la posible causa de un evento que ya sucedió. Para calcular la probabilidad necesita de tres probabilidades, entre las que necesita está la probabilidad *a priori* que sería la probabilidad de la hipótesis, la probabilidad del complemento de la probabilidad de la hipótesis y la probabilidad *a posteriori*; sin embargo, aunque el método bayesiano nos ofrezca un cálculo aritmético para tratar con las hipótesis, no es suficiente para tratar con genuinas hipótesis científicas, pues para tratarlas necesitamos exponer los problemas que se relacionan con las hipótesis científicas cuando intentamos apoyar una hipótesis dada la evidencia. Anteriormente en el capítulo tres expusimos algunos de los problemas sobre la evidencia.

Si queremos relacionar el método hipotético deductivo con el método bayesiano, de acuerdo a la interpretación de Salmon sería de la siguiente forma:

Si los elementos del método hipotético deductivo son una hipótesis *H*, afirmaciones de condiciones iniciales *I* y una predicción observacional *O* inferible, podemos decir que *H* involucra una predicción observacional *O*. Luego, si *B* se refiere a que ocurrió la predicción observacional *O*; *C* se refiere a lo que se observó con respecto a *O,* Salmon asegura que, "si ocurre *B,* significa que obtuvimos *O*; en el caso negativo '*C*' designa la falsedad de *O*."[96]

La interpretación que da Salmon se puede contrastar con el teorema de Bayes y podemos calcular la probabilidad de la hipótesis que nos interesa. Si *A* se refiere a *H* e *I*, podemos decir que en la regla de Bayes, queremos calcular la probabilidad de *A* dado *B*. Para Salmon ésta es la probabilidad que nos conduce a la corroboración de una hipótesis científica; *P* (*A*I *B*).

Por ejemplo:

H: un individuo en específico *x* que tiene la herencia *y* es de la comunidad *D*

[96] [Salmon 1984, p.117]

I: se sabe que cierta herencia *y* prevalece en la comunidad a la que pertenece el individuo *x*

O: el individuo da positivo a los experimentos

Luego, si sabemos que en dos comunidades *D* y *E* se tiene una probabilidad previa no desdeñable de que la prevalencia de la herencia *y* es *P*(4/5) en la comunidad *D* y *P*(2/5) en la comunidad *E*. Esto es, en la comunidad *D* la probabilidad de que los individuos tengan la herencia *y* es $P(H \mid D)$= 4/5, es decir, la probabilidad de que un individuo tenga la herencia *y* dado que es de la comunidad *D* es de 4/5. En la comunidad *E* la probabilidad de que los individuos tengan la herencia *y* es $P(H \mid E)$= 2/5, es decir, la probabilidad de que un individuo tenga la herencia *y* dado que es de la comunidad *E* es de 2/5. Ahora bien si queremos calcular la probabilidad de que un individuo *x* con la herencia *y* sea de la comunidad *D*, aplicamos la regla de Bayes de la siguiente manera:

$$P(D \mid H) = \frac{P(D)\, P(H \mid D)}{P(D)\, P(H \mid D) + P(E)\, P(H \mid E)}$$

La probabilidad de que el individuo pertenezca a una de las dos comunidades es de ½, pues sólo hay dos posibles comunidades que tienen individuos con la herencia *y*, a saber, *D* y *E*. Por consiguiente:

$$P(D \mid H) = \frac{P(1/2)\, P(4/5)}{P(1/2)\, P(4/5) + P(1/2)\, P(2/5)} = 4/6$$

Para calcular la probabilidad con la regla de Bayes necesitamos la probabilidad de $P(H \mid D)$, la $P(D)$ y la probabilidad de: $P(1/2)\, P(4/5) + P(1/2)\, P(2/5)$, estas son las tres probabilidades que se necesitan tener para poder aplicar la regla de Bayes.

Este resultado puede resultar erróneo, pues la probabilidad *a priori* que hemos asignado, no se ha sido del todo específico, pues puede ocurrir que cuando nos referimos a la población en general probablemente no será aplicable a un individuo *x* que consulta en un hospital y al que se le realiza la prueba, pues puede haber otros motivos de los cuales sospechamos que pueden influir de manera relevante de sospecha que no se han tomado en cuenta, por ejemplo: que el individuo pertenece a un grupo selecto genéticamente hablando o porque viene de una familia genética muy distinta a la que hay en la comunidad, datos por el estilo que hacen cambiar la probabilidad *a priori* de que el individuo *x* tenga la genética *y*. Estos datos con características cualitativas pueden consideradas por la persona que realiza el cálculo probabilístico, pero, como dice Salmon, en última instancia la decisión de tomar en cuenta estos factores como evidencia relevante depende de la experiencia de la persona que realiza el cálculo probabilístico.
Sin embargo, los cálculos son válidos si pensamos en una población en general y no hemos dado detalles como los ya citados que pueden afectar en la decisión de corroborar o de refutar una hipótesis.

En el método bayesiano se tiene la eficacia de poder cambiar la probabilidad con más datos, por ejemplo si sabemos que hay más de dos grupos de comunidades en las que sus individuos pueden tener la herencia *y*, podemos añadir esta información y, por consiguiente, nos daría otro resultado en el cálculo de la probabilidad. Podemos modificar la probabilidad conocida *(a priori)* de la hipótesis cuando tenemos nueva información al respecto. En cambio, cuando se usa el método frecuencialista de la probabilidad, se calcula la probabilidad de observar una frecuencia suponiendo que la realidad sea de una manera determinada, pues, recordemos que el método frecuencialista sólo se usa bajo los conceptos de la teoría frecuencialista de la probabilidad, es decir, está determinada bajo la teoría objetiva de la probabilidad, por ello, parte de eventos objetivos en los que no se cuestiona su procedencia (probabilidad *a priori*) que haya tenido anteriormente.

El método bayesiano sí cuestiona la probabilidad *a priori* de tal evento, por consiguiente, puede tener un mejor resultado en el cálculo de la probabilidad del evento. Sin embargo, la probabilidad *a priori* se calcula independientemente del método descrito por Salmon (método hipotético deductivo científico) y del método de probabilidad que se use, esta probabilidad *a priori* de nuestra hipótesis, origina muchos puntos de vista para calcularla, pues hay dificultades para interpretarla, por consiguiente preguntarnos por cómo debemos dar sentido a esta probabilidad no tiene una única respuesta. La interpretación que se asuma para designar probabilidades *a priori* afectará en el cálculo de la probabilidad y por ende el resultado de la probabilidad *a posteriori* (probabilidad resultante de la aplicación del método bayesiano) de las hipótesis científicas en cuestión; así pues, en la medida en que las probabilidades *a priori* de las teorías hipotético deductivas sean tomadas serán corroboradas las hipótesis. Para solucionar el problema, se pueden utilizar varios criterios, en este texto sólo mencionaremos algunos para exponer la dificultad de tratar con probabilidades *a priori*, en este caso los criterios de falsificación de Popper y de plausibilidad, podemos ver como se usan para calcular la probabilidad *a priori*.

Salmon (1984) explica estos dos criterios y expone cómo pueden ser utilizados para tratar el problema de las probabilidades *a priori*. Por un lado, Popper asegura que cuando se aplica el criterio de falsificación a las hipótesis, se considera que una hipótesis tiene más probabilidad, si es la que más se ha intentado falsificar y, no obstante, ha pasado las pruebas experimentales; por otro lado, la hipótesis que se le aplica el criterio de falsificación, es decir, se ha intentado falsificar la hipótesis con varias pruebas experimentales, y si la hipótesis es la que menos ha pasado las pruebas, entonces se considera que es la hipótesis menos probable o algunos pueden considerarla hasta falsa.

De esta forma, el deductivismo de Popper es un método de eliminación (falsificación). Una objeción a este método es que cuando una hipótesis se ve aparentemente falsificada por la evidencia, el científico puede concluir que el experimento fue errado o que las condiciones iníciales de la hipótesis son erradas. Por otra parte, el procedimiento de la inducción por eliminación no es adecuado para la inferencia eliminativa fundada en el teorema de Bayes, porque si hubiera un interminable número de hipótesis alternativas para explicar un fenómeno de ello no se sigue que todas sean plausibles.

Ahora bien, si usamos el argumento de plausibilidad en las hipótesis científicas para interpretar su probabilidad *a priori*, éste será un intento para elegir una apropiada referencia de clases de teorías plausibles; sería averiguar qué tipo de hipótesis son probables, ésta es la forma en cómo Hanson describe la plausibilidad de las hipótesis. De esta manera, las hipótesis que pertenecen a la clase de conjeturas plausibles son las que tienen alta probabilidad y las hipótesis o teorías que pertenecen a la clase de conjeturas absurdas son las que se consideran que tienen una baja probabilidad *a priori*. Para calcular la probabilidad *a priori* Salmon considera que el método frecuencialista podría ser usado de una manera eficaz, así pues, él se pregunta por cómo se determinan y qué consideraciones son pertinentes para la plausibilidad.

Salmon afirma que las características necesarias para la verdad[97] o falsedad de una hipótesis científica son propiedades que determinan una referencia de clases, por ello, para calcular la probabilidad de una hipótesis *H*, ésta debe pertenecer a una clase *A*, por ejemplo, si nuestra hipótesis es que cierto individuo *x* tiene una característica genética *y*, esta hipótesis pertenece a una clase de hipótesis que es determinada por los expertos en esta disciplina, en este caso la medicina y más específico en pruebas de genética; aun puede ser más específico. Ahora bien, hay varias formas de juzgar la plausibilidad de las hipótesis, para Salmon hay tres criterios generales que pueden ser usados como base para los juicios de la plausibilidad; estos criterios reafirman la importancia de clases, pero también pueden ser vistos como criterios de plausibilidad que las hipótesis deben confrontar. Los tres criterios son: formal, pragmático y material. El criterio formal se basa en las relaciones deductivas entre hipótesis, trata sobre las inferencias que podemos hacer cuando dos hipótesis se relacionan de una forma que pueden llevarnos a aceptar otra hipótesis. El criterio pragmático es menos fiable, pues se basa en una reacción parcialmente emocional y hasta puede ser considerada como una falacia deductiva; un ejemplo de este criterio sucede cuando descartamos la hipótesis sostenida por una persona que es religiosa y no sabe lo básico de física y, sin embargo, sostiene una teoría física. Si consideramos estos datos de la persona, de inmediato diremos que no es una teoría plausible, pues la persona no sabe nada del campo de la física; aunque parezca poco riguroso este criterio es utilizado frecuentemente por los científicos. Por último, el criterio material hace referencia a la similitud o incompatibilidad que puede existir entre las hipótesis, si las hipótesis exitosas

[97] Cuando Salmon habla de verdad o falsedad se refiere a cuando suponemos la afirmación de una hipótesis (verdad) y a la no afirmación de una hipótesis (falsedad).

tienen similitud con otras hipótesis, entonces se infiere que éstas también tendrán posibilidad de tener éxito y son consideradas como plausibles (en donde para decir que dos hipótesis son similares, se apela a sus características comunes).

Noción de probabilidad *a priori*

Salmon menciona tres criterios para calcular la probabilidad *a priori*, dos de ellas son la interpretación personalista que considera las probabilidades como grados de creencia; una hipótesis tiene cierto grado de convicción, ésta es la probabilidad *a priori* para ellos. El problema que se puede generar cuando se tienen distintas probabilidades *a priori* se cree que se soluciona cuando desaparecen las diferentes opiniones a la luz de nueva evidencia; sin embargo, esto no soluciona el problema, más adelante en el apartado a las objeciones al teorema de Bayes, se dirá por qué no se soluciona.

La segunda interpretación es la frecuencialista que define la probabilidad como el límite de la frecuencia relativa donde ocurre un atributo en una secuencia infinita de eventos. En esta interpretación hay algunas dificultades para explicar la relación entre la evidencia e hipótesis científicas, pues no pueden ser entendidas sólo en conceptos frecuencialistas, además tiene varias nociones que son cuestionables. Otro problema que tiene el método frecuencialista es saber cuándo considerar que una hipótesis científica es corroborada, este problema se origina cuando la teoría frecuencialista sólo se basa en un número de instancias corroboradas y no da una variedad de pruebas, esta objeción es la más enfatizada por Salmon y cree que es sumamente importante considerarla para la corroboración de hipótesis científicas. Una objeción relacionada con la anterior es que la teoría frecuencialista sólo se basa en instancias que pueden ser enumeradas y de este modo sólo se limita al campo cuantitativo; sin embargo, la variedad de pruebas también es un problema, pues no se tiene un criterio fiable para decir qué pruebas deben ser aplicadas. Salmon sugiere como solución tratar de realizar pruebas bajo el dominio de aplicabilidad de la hipótesis estudiada, la idea básica es hacer varias pruebas a la hipótesis; de esta manera, la variedad de instancias nos ayuda a eliminar otras hipótesis. Pero hay que tener en cuenta el campo de estudio donde se encuentra el fenómeno que deseamos explicar por medio de la hipótesis, pues desde este campo se tienen que plantear distintas pruebas.

En conclusión, la gran dificultad que tiene el teorema de Bayes es calcular la probabilidad *a priori*; Salmon dio varias maneras[98] de solucionar el problema, pero ninguna es completamente satisfactoria. De esta forma, Salmon nos dice que el resultado de la aplicación del teorema de Bayes depende de las probabilidades de sus elementos que son tres probabilidades las que necesitamos tener para aplicar el teorema, éstas ya las mencionamos en el ejemplo anterior donde aplicamos la regla de Bayes, y por supuesto una de ellos es la probabilidad *a priori*, el problema de fondo es qué debemos de considerar como probabilidad *a priori*.

Sin embargo, a pesar de las dificultades que hay en la interpretación de la probabilidad *a priori*, el método bayesiano y el método hipotético deductivo científico dan origen a una relación entre la ciencia empírica y objetiva, lo que nos permite mostrar la

---

[98] Los tres criterios que expusimos y que Salmon afirma para considerar qué hipótesis son plausibles, tienen implícito el interés de Salmon para mostrar que una interpretación frecuencialista de la probabilidad puede ser usada para la interpretación de las probabilidades *a priori*. Hemos visto que los tres criterios tienen distintas formas para decir qué clase de hipótesis son probablemente exitosas. Pero, Salmon cree que el último y primordial criterio es la experiencia; los que están de acuerdo con esta justificación de las probabilidades *a priori*, según Salmon, pertenecen al campo de los frecuencialistas. De esta forma, Salmon utiliza la probabilidad frecuencialista para tratar el problema de la probabilidad *a priori* del teorema de Bayes; sin embargo, más adelante veremos que la probabilidad frecuencialista también tiene problemas en sus criterios para calcular la probabilidad de una hipótesis científica y en general algunas de sus nociones son problemáticas.

relevancia de las probabilidades en la predicción, teoría, y decisiones prácticas; esto es una ventaja, pues la teoría frecuencialista difícilmente puede establecer tal relación. A pesar de los problemas que implica utilizar la probabilidad, complementar los métodos científicos con teoremas formales y rigurosos, como el teorema de Bayes nos lleva cada vez más acercarnos a un resultado más estable, aunque a veces tenemos que ceder o perder un poco de formalismo y rigor, por ello, la siguiente pregunta sería, qué estamos dispuestos a ceder para obtener más estabilidad.

El método de Bayes por Jeffreys

En el capítulo 1 en la sección: el problema de Bayes y su solución mencionamos el caso de Harold Jeffreys (1931) que sostiene una filosofía de la probabilidad, como investigador en su campo de estudio, afirma que constantemente se enfrentaba a problemas de probabilidad inversa y usaba como método la regla de Bayes que consideraba como una herramienta necesaria en los métodos científicos que empleaba para corroborar o rechazar hipótesis científicas, Jeffreys decía: "[…] el teorema de Bayes es a la teoría de la probabilidad, como el teorema de Pitágoras es a la geometría."[99]

Jeffreys frecuentemente utilizaba el teorema de Bayes como método de investigación científica. Si bien trabajaba en el campo de la geofísica, Jeffreys asegura que constantemente no disponía de suficientes datos para rechazar varias hipótesis. La teoría que utilizaba, que era el método más reconocido por los contemporáneos, la teoría frecuencialista, no era muy eficiente para él, pues en su experiencia como científico, afirmaba que constantemente se enfrentaba a problemas de probabilidad inversa. Como dice Galavotti:

[99] [Jeffreys en Galavotti 2005, p.179]

> [Jeffreys] tenía que explicar datos experimentales por significativas hipótesis alternativas, o para evaluar hipótesis generales a la luz de datos que cambiaban constantemente, en ese momento Jeffreys comenzó a trabajar sobre esta clase problemas usando el método bayesiano, sin embargo como decíamos en aquella época el método bayesiano no era apreciado por los científicos y estadísticos que se aferraban al frecuentismo para usar este método en sus investigaciones científicas. [Galavotti 2005, p.180]

De esta forma, Jeffreys consideraba que la probabilidad es la guía más fundamental y general de la ciencia, además de que la probabilidad representa una rama de la ciencia.

Galavotti nos muestra el inductivismo de Jeffreys que parte desde una visión epistemológica de la probabilidad, pues considera que hay una relación entre una proposición y un conjunto de datos que expresan un grado de creencia racional, es decir, dado un conjunto de datos, una proposición q tiene, en relación a estos datos, una y sólo una probabilidad; pero, si una persona asigna una probabilidad diferente, se dirá que esa persona está equivocada. Bajo este supuesto, Galavotti dice que Jeffreys está convencido de que una teoría satisfactoria tiene sólo una probabilidad y se relacionan con la convicción de que sólo existe un grado de creencia racional para tal teoría. Hemos de notar que la noción de creencia racional se refiere a la interpretación lógica de la probabilidad y difiere del subjetivismo, como ya mencionamos en el capítulo dos, la diferencia de estas teorías de la probabilidad radica en que la teoría subjetiva de la probabilidad considera los grados de credibilidad del individuo que realiza el cálculo de probabilidad, en cambio la interpretación lógica de la probabilidad de Keynes afirma que todos tenemos una misma intuición lógica (grados de credibilidad racional) y por ende podemos llegar a un acuerdo, al inferir de la misma forma por medio de la intuición lógica, aunque tiene el problema de justificar qué entiende por intuición lógica. Otro punto importante que menciona Jeffreys es el trato de la interpretación subjetiva de la probabilidad, pues este tipo de interpretación se consideró como erróneo, ya que sólo se creía que la interpretación objetiva podía ofrecer rigor; sin embargo, Galavotti afirma que no hay razones para rechazar la interpretación subjetiva, pues es tan válida como cualquier interpretación de la probabilidad, por ello, está en desacuerdo con Jeffreys cuando no utiliza la teoría subjetiva de la probabilidad, pero si la regla de Bayes. Sin embargo, Galavotti no explica cómo Jeffreys calcula la probabilidad a priori. Si bien no apela al método frecuencialista, como lo hace Salmon, entonces es posible que se funde en su experiencia en su campo de estudio para considerar qué hipótesis es más plausible.

De esta forma, la regla de Bayes, independientemente de bajo qué interpretación de la probabilidad se use, nos proporciona una forma de calcular la probabilidad de hipótesis a la luz de evidencia observable, aunque con el problema de cómo definir la probabilidad *a priori*. Hay quienes afirman que la definición de probabilidad *a priori* depende de criterios objetivos; Salmon es uno de ellos y, por ello, considera la probabilidad frecuencialista como un método alternativo para calcular la probabilidad *a priori*, pero también hay subjetivistas como Galavotti que toman una actitud menos formal y, como habíamos dicho, el problema está relacionado con el problema que tienen todas las teorías de probabilidad que es su interpretación de la probabilidad y en el fondo el problema es carecer de un concepto único de la probabilidad. Sin embargo, puede haber personalidades (como Howson y Urbach[100]) que no creen que el cálculo de la probabilidad *a priori* sea problemático, pues es similar al hecho de que en la lógica deductiva nos interese la validez de un argumento y no la verdad de sus premisas. Así pues, para Howson y Urbach las inferencias bayesianas serian válidas o no, independientemente de qué tan acertada es la

[100] [Colin Howson y Peter Urbach 2006, pp.13-114]

probabilidad *a priori* asignada inicialmente. El problema es que en los métodos científicos para corroborar hipótesis científicas no es suficiente con que la inferencia sea válida, pues el cálculo de la probabilidad *a priori* puede afectar de manera considerable la probabilidad de una hipótesis y, por ende, la decisión de aceptarla o rechazarla.

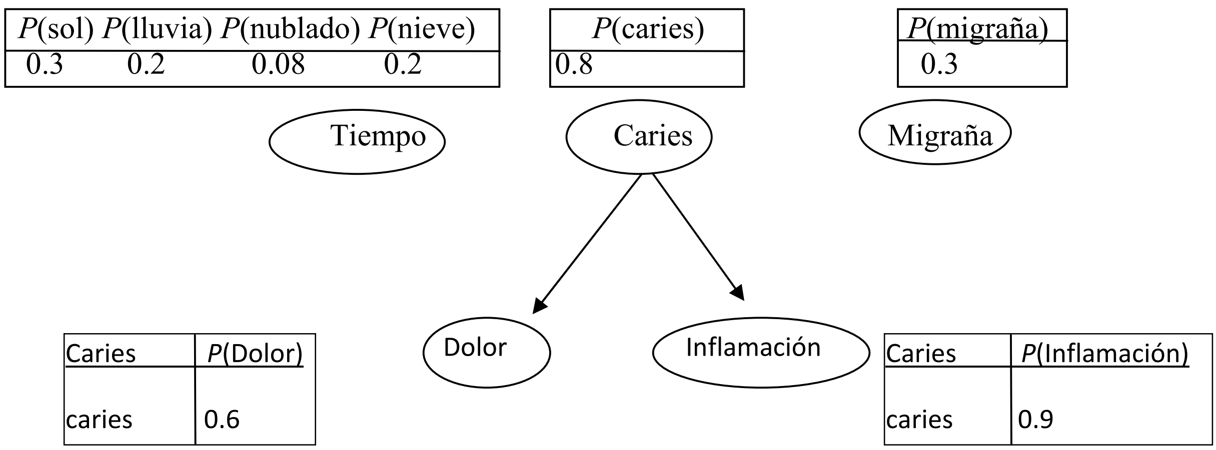
Las inferencias que realizamos cotidianamente o en investigaciones científicas están de acuerdo con la estructura de la regla de Bayes. Por ejemplo, cuando un dentista hace inferencias para diagnosticar a un paciente, por lo general, se guía por los síntomas que observa en el paciente y de ahí infiere qué tiene el paciente para darle una medicina, es este tipo de estructura causal que constantemente usamos para explicar los hechos, y esta estructura la usa la regla de Bayes para calcular la probabilidad.

También las inferencias que un detective realiza tienen la estructura que usa la regla de Bayes, como muestra Galavotti en su ejemplo:

> De esta manera, cuando un cuerpo es encontrado, es virtualmente cierto que un asesinato ha sido cometido, y una huella digital en la manija de una pistola puede ser considerada como una evidencia determinante para suponer que una cierta persona *X* ha disparado. Incluso en estos casos de inferencia tiene la estructura de la regla de Bayes y es muy frecuente ver en estos casos las valoraciones que se hacen de las probabilidades del antecedente. [Galavotti 2005, p.51]

Este ejemplo, también muestra que en la investigación que realiza un detective para inferir los motivos del criminal es un intento para calcular la probabilidad del antecedente del caso, de una misma forma, la inferencia inductiva para la observación de datos que pueden corroborar una teoría científica; en este caso podemos observar que tiene la estructura de causa y efecto, donde el efecto son los datos y la posible causa son las hipótesis; bajo esta estructura se usa la regla de Bayes. En cualquier ejemplo nos muestra que la forma causal que usamos para encontrar las causas de hechos, observamos los hechos e intentamos dar explicaciones de su causa; ésta es la forma que usa la regla de Bayes para calcular la probabilidad de una posible causa (hipótesis).

Otro campo de investigación en donde se aplica la regla de Bayes es en inteligencia artificial. La regla de Bayes es usada en redes bayesianas; éstas pueden ser utilizadas para predecir, clasificar o diagnosticar, entre otras aplicaciones. Una red bayesiana tiene la característica de permitir estudiar las relaciones de dependencia y causalidad, constituye una manera práctica y simple de representar el conocimiento incierto. Por ejemplo, queremos calcular la probabilidad de posibles hipótesis que expliquen los síntomas de un paciente. Sabemos que el paciente tiene los síntomas de dolor muscular, tos y vómito; desde esta evidencia, los médicos, con base en su experiencia pueden tener varias hipótesis alternativas, y dada la evidencia se calcula la probabilidad de cada una de ellas. En este sencillo ejemplo se puede utilizar la regla de Bayes para calcular la probabilidad de cada hipótesis considerada, dada la evidencia. Otro ejemplo del uso de la regla de Bayes en las redes bayesianas puede ser graficado: supongamos que nos duelen los dientes, luego vamos al dentista, el dentista observa los síntomas que son dolor e inflamación en la boca; conforme a la experiencia del doctor y datos probabilísticos que tiene, puede descartar varias hipótesis que expliquen los eventos, podemos graficar esto de la siguiente forma:



En esta red bayesiana, podemos ver algunas posibles hipótesis que pueden explicar el evento observado, en este caso los síntomas; usando la regla de Bayes se ha calculado la probabilidad de cada hipótesis considerando la evidencia; al igual que la probabilidad de que sea caries dado que hay dolor e inflamación. Intuitivamente la caries está relacionada, por lo general, con el dolor e inflamación en la boca, y la probabilidad de cada hipótesis nos demuestra de una manera gráfica que la hipótesis con más probabilidad es la hipótesis de que la caries sea la causa. Aunque el ejemplo es sencillo, este tipo de inferencia científica es muy usada en casos complejos donde no sólo se tienen tres hipótesis y el caso muy conocido; además intuitivamente podemos saber qué hipótesis tiene más probabilidad. Sin embargo hay casos en donde es muy difícil saber cuál es la hipótesis más probable si se tienen varias hipótesis, pues cuando se rechazan, pueden surgir otras. Por otro lado, cuando las probabilidades de cada hipótesis son muy semejantes o el caso en que si los datos son muy pocos, varias hipótesis son muy probables. A pesar del uso del teorema de Bayes en la corroboración de hipótesis científicas, tiene algunas deficiencias, que enseguida abordaremos.

4.1.2 Deficiencias y eficiencias sobre el uso del método bayesiano en la corroboración de hipótesis científicas

Recordemos que Gillies al final del apartado: “teoría subjetiva” le señala al teorema de Bayes en general algunas objeciones, unas van más dirigidas a la interpretación subjetiva del teorema de Bayes. En el primer problema que Gillies atribuye al teorema de Bayes ocurre cuando tenemos que calcular la probabilidad de una hipótesis dada cierta evidencia y para hacerlo utilizamos condicionalización bayesiana; el problema surge cuando tenemos que considerar una secuencia de eventos para calcular la probabilidad de la hipótesis, pues podemos tener una secuencia de eventos, pero no sabemos qué tan larga tiene que ser la secuencia; por ejemplo, para calcular la probabilidad de que mañana llueva, depende sobre los días previos, pero no sobre los días de hace 100 años, es claro que esta clase de dependencia tendrá una secuencia larga, pero no sabemos hasta qué punto la secuencia será muy larga y de este modo nos pueda alejar del valor correcto, o podría ser demasiado corta lo que tampoco nos permitiría acercarnos al valor correcto[101].

Otra objeción sucede cuando aprendemos de cierta experiencia A y nuestro juicio influenciado por la experiencia puede inclinarse por la probabilidad de que suceda un evento E, pero $P(E_{n+1}$ I $A)$, es decir, si nuestra opinión inicial E, nuestra probabilidad *a priori*, es rechazada a la luz de nueva evidencia, bien podemos cambiar la evidencia o esperar nueva evidencia para seguir fieles a nuestro evento E. De este modo se cuestiona el supuesto de de Finetti quien trata de responder cómo llegan a tener una misma probabilidad *a posteriori* sobre cierto evento distintos individuos; su solución, según Gillies, es que a la luz de nueva evidencia pronto llegarán a converger sus opiniones. Sin embargo, no sabemos cuándo sucederá; por ejemplo, no sabemos si una persona quedará convencida a la luz de la evidencia que arrojan 500 eventos o si quedará convencida con la evidencia que arrojen 700 eventos, es decir, después de 700 eventos sus probabilidades *a posteriori* pueden seguir siendo totalmente diferentes. No tenemos garantía cuando sucederá tal convergencia.

---

[101] Hemos de notar que Gillies como objetivista de la probabilidad supone que hay un valor correcto o exacto, y cuando aplicamos probabilidad el resultado que siempre nos dará es sólo una aproximación al valor exacto que supone Gillies. La noción de valor correcto es objetable pues no está justificado este supuesto, no hay razones para creer que hay un valor exacto que nunca alcanzaremos. Además, el problema que menciona Gillies sobre las secuencias demasiado grandes o pequeñas es también atribuido a la teoría frecuencialista de la probabilidad que está bajo la interpretación objetiva de la probabilidad.

Ahora bien, la noción de intercambiabilidad que de Finetti introduce para explicar la probabilidad de un evento *A* dada una evidencia *B*, es igual a la probabilidad inicial de *A*, es decir, de que suceda *B* o no *B*, no afecta la probabilidad de *A*, esta definición es la misma que la noción objetiva de 'independencia'[102], que es equivalente a la noción subjetiva 'intercambiabilidad', pues recordemos que expresan lo mismo. Por ello, la noción subjetiva de intercambiabilidad es plausible cuando se aplica a eventos como los volados, donde cada volado que se realiza es independiente del anterior; el problema, según Gillies, sucede cuando la noción de intercambiabilidad no es adecuada para los eventos que son dependientes[103]. Si no tenemos garantía de que los eventos son independientes, el uso de intercambiabilidad puede dar resultados poco acordes con la realidad.

Gillies ejemplifica el problema simbolizando el fenómeno de que salga el sol con un experimento aleatorio donde tenemos una caja que contiene bolas amarillas y negras; de acuerdo a nuestra experiencia hay 1000 bolas amarillas y puede haber dos bolas negras que expresen la no salida del sol, en este tipo de eventos sabemos que la selección de las bolas con reemplazo son independientes, por ello, Gillies dice que la noción de independencia bien se aplica a este tipo de eventos; pero si nosotros tratamos de representar las salidas del sol con este tipo de experimentos, obtendremos resultados poco coherentes con la realidad, pues las 2 bolas negras no representan ningún evento de la realidad. Gillies afirma que ante tal situación de querer representar eventos de la naturaleza con experimentos, elegir bolas desde un contenedor, es un caso absurdo en el caso donde no sale el sol, es decir, que

---

[102] Se dice que *A* y *B* son independientes si $P\ (A/\ B)$ es igual a $P(A)$ y $P\ (B/\ A)$ es igual a $P(B)$. intuitivamente podemos notar que la ocurrencia o no de alguno de los eventos no afecta la probabilidad del otro.

[103] Se dice que *A*y *B* son dependientes si $P\ (A/\ B)$ no es igual a $P(A)$ y $P\ (B/\ A)$ no es igual a $P(B)$. intuitivamente podemos notar que la ocurrencia de alguno de los eventos afecta la probabilidad del otro.

obtengamos de la urna una bola negra, lo que representaría la no salida del sol. Esto es, porque sabemos que las independencias se aplican en el caso de seleccionar bolas y esto no aplica en el caso de que salga o no salga el sol que es un fenómeno de la naturaleza[104]. Esto refuerza su convicción de que sólo podemos aplicar la noción de intercambiabilidad de de Finetti si estamos seguros sobre las bases de nuestro conocimiento que los eventos son objetivamente independientes; de aquí que Gillies también considere que es necesario que los eventos sean objetivos.

Ahora bien, según Gillies, de Finetti interpretó algunos de sus resultados matemáticos sobre la relación entre independencia e intercambiabilidad como mostrando cómo se pueden eliminar las nociones de probabilidad objetiva e independencia a favor de la noción de probabilidad subjetiva e intercambiabilidad; ésta es la reducción de intercambiabilidad de de Finetti que Gillies objeta, pues los resultados matemáticos prueban lo opuesto: sólo se puede aplicar intercambiabilidad cuando la situación es objetiva e independiente, por ello, para Gillies el concepto de intercambiabilidad es inútil y de Finetti, según Gillies, inicia desde la noción objetiva de independencia para poder aplicar intercambiabilidad. En este sentido podemos decir que un evento es independiente si asumimos el método frecuencialista de la probabilidad, pero si asumimos el método bayesiano de la probabilidad podemos asumir que un evento es independiente o un evento es intercambiable, depende de qué teoría de la probabilidad asumamos, así pues, podemos ver que el uso del método bayesiano tiene la facilidad de poder tratar con eventos independientes o intercambiables.

---

[104] Sin embargo también podemos decir que como fenómeno de la naturaleza no podemos tener una secuencia exacta, sólo podemos tener una secuencia si nosotros la creamos a partir de nuestra interpretación de los eventos.

En este sentido Gillies extiende su crítica al reduccionismo de de Finetti y al bayesianismo tratando de mostrar que su atractivo principal, el cual es ofrecer una sencilla fórmula matemática para demostrar la forma en la cual una persona racional debe cambiar sus creencias a la luz de la evidencia, se pierde cuando se toma conciencia de que, aunque la gente racional puede cambiar sus creencias de la forma en la que el bayesianismo lo sugiere, también lo puede hacer de una forma totalmente diferente.

De esta manera, Gillies crítica el enfoque bayesiano de de Finetti, cree que lleva a tantas complicaciones, pues en este enfoque, según Gillies, es necesario considerar todas las posibilidades que pueden surgir en el comienzo de la investigación, es decir, a todas las probabilidades *a priori* que distintas personas pueden tener sobre la investigación. Gillies cree que esta complicación es eliminada completamente adoptando la aproximación objetiva y sólo necesita considerar una posibilidad, es decir, en términos subjetivistas una probabilidad *a priori*. Por ejemplo, si es un popperiano objetivista, su tema inicial es una serie de pruebas estadísticas; quizás estas pruebas corroborarán su conjetura inicial, pero también la pueden rechazar, en el último caso a la luz de nueva evidencia puede intentar elaborar una nueva hipótesis; mientras que el bayesiano intenta considerar todos los posibles *a priori* y esto es totalmente irrealizable, pues de forma intuitiva sabemos que nunca se puede realizar tal hazaña, si cada vez aparecen más hipótesis alternativas y/o ignoramos la existencia de algunas hipótesis.

Otro problema que se refiere a la secuencia de eventos; por ejemplo: si tengo un conjunto de hipótesis, es decir, $\{H_\theta\}$ donde $\theta \in I$, para algún conjunto I, el parámetro θ se da por una distribución *a priori* $\mu(\theta)$ y esto es cambiado por una distribución *a posteriori*

$\mu(\theta \mid e)$ por condicionalización bayesiana. Las distribuciones se realizan sobre el conjunto de hipótesis consideradas. Por lo tanto, tenemos: $P(H_\theta)=\mu(\theta)$, la probabilidad de H_θ es dada por todos los posibles *a priori* y $P(H_\theta \mid \varepsilon) = \mu(\theta \mid e)$, la probabilidad de $P(H_\theta \mid e)$ es dada por la evidencia observada.

Gillies usa el siguiente ejemplo, para expresar el problema:

> supongamos que hay una caja negra la cual el señor *B* ve que aparece una cifra 1 o 0 en una pantalla en intervalos regulares t= 0, 1, 2, …, n, … si la secuencia de cifras es X_0, X_1, X_2, …, X_n, […] y esta secuencia es generada por algún proceso desconocido por el señor *B* quien tiene que asignar probabilidades de la forma $P(X_n \mid X_1, X_2, \ldots, X_{n-1})$, entonces puede utilizar esta secuencia para calcular la probabilidad de X_n. De esta forma él puede saber los valores de X_0, X_1, X_2, …,X_{n-1}, pero no de X_n, cuando tiene estas probabilidades las puede tomar como sus ecuaciones de apuesta en un juego de azar con la señora *A* sobre el valor de las *n* cifras. Ahora bien, si *e* es el estado del valor observado de X_0, X_1, X_2, …, X_{n-1}, el usa $P(H_\theta \mid e)$ para calcular $P(X_n \mid e)$. [Gillies 2000, p.83]

En este sistema, podemos volver a decir que la objeción es que el señor *B* elige la secuencia H_θ, θ está, $0 \leq \theta \leq 1$. Luego, podemos ver cómo el sistema de condicionalización bayesiana que utiliza el señor *B*, como su medio de aprendizaje, produce una secuencia de probabilidades completamente distintas a la realidad, pues utiliza condicionalización bayesiana para eventos que son independientes. Recordemos que la condicionalización bayesiana no tiene problemas para calcular la probabilidad de eventos dependientes; los problemas, o mejor dicho, las objeciones que Gillies atribuye a la condicionalización bayesiana sucede cuando trata de calcular la probabilidad de eventos independientes con condicionalización bayesiana, pues su uso puede dar como resultado una probabilidad que está muy alejada de la realidad; por ejemplo, calcular la probabilidad de que un individuo en específico *x* tiene la cierta herencia genética *y*; se calcula su probabilidad y el resultado que se puede obtener puede resultar intuitivamente erróneo para los expertos en el área que tratan este tema.

Por lo tanto, la condicionalización bayesiana no será considerada como una buena estrategia de conocimiento, pues H_θ es posible que no sea el X_n, es decir, la secuencia que elige el señor *B* lo hace sin ningún criterio, sólo la elige porque es la única que tiene; también, tampoco estamos seguros de que la secuencia tiene la característica de eventos independientes o dependientes.

Por último, las críticas de Gillies hacen que las características del bayesianismo se pierdan y ya no sea una teoría útil; concluye que la noción de intercambiabilidad de De Finetti no funciona y que sus argumentos contra esta reducción de intercambiabilidad pueden ser usados contra el bayesianismo en general.

Ahora bien, considero que las críticas de Gillies van más dirigidas a la interpretación de la regla de Bayes (teoría subjetiva de la probabilidad) que al cálculo probabilístico del método bayesiano. Este tipo de objeciones aún no han sido totalmente definidas como problemas de la probabilidad sino como problemas de la estadística, es decir, de la interpretación de los eventos, lamentablemente no hay un criterio que especifique si esto es un problema de la probabilidad o de otra disciplina ya sea de estadística o de epistemología.

Siendo así como he dicho antes, creo que son buenas críticas, si están bien dirigidas, pero aun así considero que el teorema de Bayes tiene muchas utilidades, siempre y cuando, se conozca en dónde y cómo es usado; por algo ha sido un teorema muy eficaz en las disciplinas científicas en la corroboración de hipótesis científicas.

Galavotti está a favor de la interpretación subjetiva del método bayesiano de la probabilidad. Ahora bien, la interpretación de la teoría subjetiva de de Finetti de la probabilidad que hemos expuesto es una interpretación de Gillies quien, recordemos, él mismo dice que es desde un objetivista hasta un pluralista de la probabilidad, por ello, encontramos varias críticas y objeciones que atribuye a la teoría subjetiva de la probabilidad. Pero ahora veremos un punto de vista subjetivo, es decir, un filósofo que es un subjetivista de la probabilidad.

El probabilismo de de Finetti es totalmente bayesiano, pues es esencial para utilizar su noción de 'intercambiabilidad'. Recordando la crítica de Gillies sobre la probabilidad inicial que cambia a la probabilidad final, Galavotti no cree que de Finetti se haya referido a un cambio de probabilidad *a priori* a probabilidad *a posteriori* a la luz de nueva evidencia, esto no significa que la persona cambie de opinión:

> Si razonamos de acuerdo al teorema de Bayes no cambiamos de opinión. Tenemos la misma opinión y nos actualizamos a la nueva situación. Si ayer yo dije 'hoy es miércoles', hoy yo digo 'es jueves'. No tengo que cambiar mi mentalidad, para el día siguiente de miércoles es en efecto jueves. [de Finetti en Galavotti 2005, p.215]

De esta forma se resuelven las objeciones de Gillies sobre el tránsito del cambio de opinión inicial a opinión final. La idea de corregir previas opiniones en la interpretación de Galavotti sobre la probabilidad es completamente extraña, en cambio considera que es un procedimiento autocorregido. Por ejemplo, siguiendo el primer ejemplo de este capítulo mostramos el cálculo de probabilidad de una hipótesis sobre el individuo x que tiene ciertos rasgos genéticos y, hicimos notar que la probabilidad obtenida sobre la hipótesis dada la evidencia era posible que cambiara con más información. En este caso, la opinión que tenían al principio puede interpretarse como una opinión que no cambia sino que sólo se actualiza con más información, pues, en este caso, si hubieran sabido que el individuo *x* al que se le realizan las pruebas era de otra comunidad no reconocida, su opinión hubiera cambiado bastante, pues ya no verían el resultado de la probabilidad como un resultado muy probable, sino que cambiaría su opinión con nueva información que consideran relevante en el cálculo de la probabilidad de la hipótesis dada la evidencia.

Una diferencia entre la teoría subjetiva y la teoría objetiva de la probabilidad, es que la teoría subjetiva admite el uso de la probabilidad *a priori*, en cambio, la teoría objetiva rechaza el uso de la probabilidad *a priori*. Para Galavotti rechazar la probabilidad inicial es rechazar el estado de la información que depende de esta probabilidad; si se rechaza, se distorsiona la probabilidad y se supone que es una entidad objetiva. De aquí la afirmación de de Finetti de que la probabilidad objetiva es una ilusión, para él es una ilusión considerar que existe una probabilidad independientemente de nosotros y del juicio de cualquier persona. Ahora bien, von Mises no cree en la afirmación de de Finetti, para él la probabilidad es independiente de quien la calcula. Otra distinción que hay entre estas dos posturas, es la distinción entre la definición de probabilidad y evaluación de la probabilidad, Galavotti cree que es una ventaja que tiene la teoría subjetiva porque hace esta distinción, en cambio la probabilidad objetiva no.

El rechazo de la interpretación objetiva de la probabilidad por parte de de Finetti no es total, pues afirma que ningún elemento subjetivista elimina los elementos objetivos y supone que una evaluación de la probabilidad expresa algún grado de creencia; además asegura que cualquier evaluación de la probabilidad esencialmente depende de dos componentes: el primero es objetivo y consiste de la evidencia de datos y hechos que conocemos; el segundo componente es subjetivo y consiste de la credibilidad con los hechos desconocidos basados sobre la evidencia conocida.

Usar la teoría bayesiana en la corroboración de hipótesis científicas tiene la ventaja de considerar la experiencia de los científicos al tomar en cuenta la influencia de su experiencia en el cálculo de la probabilidad de hipótesis.

Otra ventaja en la investigación científica de la aproximación bayesiana puede observarse cuando considera el progreso como un proceso continuo; por ejemplo, si un científico es innovador y tiene creencias *a priori* para apoyar su hipótesis, tendrá que exponer sus datos y demostrar si los datos tienen coherencia en la dirección en que lo afirma, sólo entonces las probabilidades *a posteriori* serán un poco más altas de lo que eran originalmente, pues ya tendrá apoyo de una probabilidad *a priori*, tendrá datos como antecedentes que pueden ser considerados como evidencia para la hipótesis científica, de esta forma los datos que tiene como antecedentes sería como un estado de conocimiento orientado de alguna manera como evidencia de ese planteamiento novedoso.

En conclusión, la interpretación de Galavotti sobre la teoría subjetiva de la probabilidad de de Finetti respeta la evaluación o cálculo de la probabilidad como un procedimiento formal y que es coherente con el resultado de la observación de los factores; de ahí la importancia de resaltar la recolección de evidencia y el factor subjetivo que influye en la recolección de evidencia. Sin embargo, de Finetti también tiene objeciones sobre el frecuentismo que veremos en el apartado: “eficiencia y deficiencias del método frecuencialista”.

4.2 El método frecuencialista aplicado a la teoría de la corroboración de hipótesis científicas

4.2.1 El análisis frecuencialista en la corroboración de hipótesis científicas

Salmon (1973) clasifica las hipótesis científicas en dos clases: las hipótesis que son generalizaciones universales y las hipótesis que son generalizaciones estadísticas, como ejemplo expone dos casos distintos de inferencia:

> Según la ley de Hooke, la fuerza necesaria para producir una distorsión en un objeto elástico (tal como un resorte de acero) es directamente proporcional a la cantidad de la distorsión. Se ha observado que un determinado resorte se alarga en una pulgada cuando se le aplica una fuerza de cinco libras. Se aplica ahora una fuerza de diez libras. Síguese que el resorte se alargará en dos pulgadas. Si esto es efectivamente así, el resultado constituye un caso confirmador de la ley de Hooke.
>
> Si con una moneda honesta se echan reiterados volados, las caras y las cruces ocurren al azar; pero con igual frecuencia a la larga. Puede demostrase que existe un probabilidad de 0.95 en el sentido de que 100 volados de la moneda en cuestión darán entre 40 y 60 caras. Consideremos la hipótesis de que la moneda es honesta (no trucada). Se efectúan diversos experimentos con ella, consistiendo cada experimento en 100 volados. En cada caso, el número de caras queda entre 40 y 60. En condiciones adecuadas, estos resultados confirmarían la hipótesis de que la moneda es honesta. [Salmon 1973, p.119]

En estos ejemplos Salmon quiere mostrar la distinción entre hipótesis y, por ello, la distinta forma de proceder para tratarlas. El ejemplo *a)* tiene la forma de causa y efecto, esta forma nos permite deducir que el resorte experimentará un alargamiento de dos pulgadas; así hemos asignado a la hipótesis como la causa y el evento como el efecto, de esta manera, estamos exponiendo la forma en que una hipótesis o mejor dicho varias hipótesis alternativas intentan explicar un evento que ya sucedió, en el método científico se intenta corroborar la hipótesis.

En cambio, el ejemplo *b)* tiene forma inductiva, pues con base en frecuencias podemos calcular la probabilidad de la hipótesis que se refiere a generalizaciones estadísticas; para Salmon esta es la forma del silogismo estadístico. En este caso el método frecuencialista puede ser un método muy eficaz para calcular la probabilidad de la hipótesis.

Ahora bien, el método frecuencialista ha sido caracterizado por Gillies como un método que tiene problemas para tratar eventos cualitativos, pues sale de su campo de estudio, en cambio otras teorías de la probabilidad pueden abarcar los eventos cualitativos. Para eventos cuantitativos Gillies cree que es mejor utilizar el método frecuencialista. También el propio von Mises considera que su método frecuencialista es más eficaz en los casos estadísticos. Pero, cuando se usa el método hipotético deductivo científico para corroborar o rechazar hipótesis y la teoría frecuencialista, ésta última a veces no es muy útil para calcular la probabilidad de las hipótesis alternativas, pues cuando se tratan tanto eventos cuantitativos como cualitativos, sólo puede manejar los eventos cuantitativos, por consiguiente, se opta por usar un método que pueda manejar ambos tipo de eventos. De esta forma la teoría frecuencialista como método de corroboración de hipótesis científicas, bajo tal supuesto no es efectiva.

La probabilidad frecuencialista puede ser aplicada en el ámbito de la investigación científica como un método para asignar probabilidad a hipótesis alternativas, bajo las definiciones que afirma la teoría frecuencialista: colectivos, clases de eventos, secuencia de eventos, frecuencia límite y las leyes que expone, como la ley de estabilidad de frecuencias estadísticas.

Sin embargo, la aplicación de la teoría frecuencialista en la teoría de la corroboración de hipótesis científicas es muy poco usual, pues las varias restricciones que von Mises expone para poder usar dicha teoría hacen que la teoría sea poco accesible en su uso. Por ejemplo, si recordamos el primer ejemplo de este capítulo que refiere a la hipótesis: un individuo en específico *x* tiene la herencia genética *y*, las pruebas que se realizan para corroborar la hipótesis son de distintas clases que pueden tratar con evidencia que es no sólo cuantitativa (se sabe que la población de la que proviene tal individuo hay una alta probabilidad de que las personas tengan la herencia genética *y*) pero también puede haber evidencia con características cualitativas (el individuo no nació en esa población, pero fue registrado como si hubiera nacido ahí).

Para aplicar la teoría frecuencialista se debe de hablar de colectivos de un cuerpo de evidencia y de un atributo a destacar '*A*' que a veces sucede en una secuencia y a veces no, es decir, es aleatorio. Sin embargo la mayoría de los fenómenos a explicar por medio de hipótesis muy pocas veces pueden acatarse a las definiciones necesarias para calcular la probabilidad; además puede haber eventos que no tienen una secuencia ordenada y, por ello, no pueden ser calculados en la teoría frecuencialista; pueden ser eventos que no se repiten. Por ejemplo, en el ejemplo mostrado sobre el individuo en específico *x* que tiene una herencia genética *y,* queremos saber si pertenece a cierta comunidad A, debemos aplicar pruebas, pero una de ellas puede afectar a los resultados que podamos obtener de las demás pruebas que realicemos al individuo *x* porque los resultados que obtengamos pueden estar a favor de la hipótesis y otras en contra, es decir, podemos obtener resultados (evidencia) que se a favor del rechazo de la hipótesis, en este caso son pruebas las que se realizan y no instancias, por ello, no puede ser una secuencia ordenada de las pruebas que se realicen, pues la prueba que afecta a las demás puede realizarse al principio y afectar la probabilidad; por otro lado, estamos hablando de pruebas y no de ejemplificaciones, por consiguiente, no podemos decir que tenemos una secuencia de ejemplificaciones sino pruebas que tienen ejemplificaciones, por esto, no podemos aplicar la probabilidad frecuencialista, pues sólo se aplica a ejemplificaciones, es decir, a eventos que son instancias de una prueba o mejor dicho de un experimento. Ahora bien, las pruebas que se realicen se deben relacionar entre ellas las probabilidades que arrojen de su aplicación y esto no puede hacerlo el método frecuencialista. Pero lo que si puede hacer la teoría frecuencialista es que todas las probabilidades bajo la noción de frecuencia límite pueden ser condicionales, pues dependen de un colectivo y no de un grado de creencia, esto puede ser visto como una ventaja para el método frecuencialista, pues si queremos calcular la probabilidad de las hipótesis alternativas sería conveniente que fuera lo más rigurosa y formal posible, para evitar prejuicios y hasta intenciones maliciosas que pueden afectar la probabilidad de una hipótesis que puede ser una buena alternativa para explicar tal evento.

4.2.2 Deficiencias y eficiencias sobre el uso del método frecuencialista en la corroboración de hipótesis científicas

Si bien el método frecuencialista para calcular la probabilidad de hipótesis científicas alternativas es rara vez usado en el método hipotético deductivo científico, esto se debe a las restricciones que von Mises atribuye a la teoría frecuencialista para obtener un método riguroso, objetivo y formal; para obtener su método tuvo que ceder la ventaja de poder aplicar la teoría frecuencialista a eventos cualitativos que, por lo general, están relacionados con la investigación científica y en la vida real.

Un problema de la teoría frecuencialista es usar el concepto de clases, si bien recordamos este término es importante en la interpretación frecuencialista de von Mises, el problema es seleccionar la apropiada referencia de cada clase, pues para algunos pueden ser relevantes tales atributos y para otros no, es decir, no hay un criterio formal para definir las clases que han de usarse para calcular la probabilidad de un evento. Pero Reichenbach, afirma que se debe elegir el atributo que dé una referencia más restrictiva de clases para las estadísticas fiables, de esta forma se trata de evitar el problema de seleccionar la clase correcta; en cambio Salmon diría que una solución sería:

> Debemos referirnos a la más amplia clase de referencia homogénea de la que es miembro ese susodicho. En cualquiera de las formulaciones, el intento es bastante sencillo. Una probabilidad debe ser establecida inductivamente y con el fin de determinar la probabilidad que los casos deben ser suficientes para poder hacer una generalización inductiva. [Salmon 1984, p.91]

De esta forma quiere evitar referirse a clases que son demasiado restrictivas, pues dice que cuando se hacen las clases restrictivas se hace que se carezca de suficientes pruebas para fundamentar una inferencia; también afirma que las clases deben de contener casos que son relevantes y que están relacionados. Cuando Salmon habla de casos relevantes se refiere a que debemos de tener relevancia estadística, para él este es el concepto clave en la solución del problema de las clases.

Otra deficiencia o crítica que se atribuye a la teoría frecuencialista en general es cómo se especifican los colectivos, no hay un punto de partida que nos guíe, por ello, hay muchas formas en que se puede definir un colectivo empírico; como habíamos visto en la interpretación objetiva de la probabilidad, un colectivo empírico no tiene una secuencia ordenada, sino que nosotros tenemos que dársela bajo nuestros criterios, esto implicaría que cada quien podría tener una distinta secuencia y consecuentemente también afectaría la probabilidad del evento que deseamos calcular.

Ahora bien, en la teoría frecuencialista de von Mises las secuencias infinitas no tienen una representación en la realidad, pues las secuencias sólo tienen eventos observables, y si los eventos son observables, las secuencias deben ser finitas, por ello, no hay secuencias infinitas observables, intuitivamente esto es cierto; por ejemplo: no podemos lanzar una moneda infinitamente. El problema de fondo es la relación que quiere establecer von Mises entre teoría y observación. Si bien la frecuencia límite de la definición de probabilidad es un concepto teórico matemático, en la teoría frecuencialista este límite depende de un concepto observable, este es la frecuencia finita. Esta es una crítica a la teoría de von Mises, pues su solución para relacionar los eventos empíricos con su cálculo formal de la probabilidad es discutible.

Una situación que refleja el mismo problema está dada por los colectivos empíricos finitos. El problema es si pueden ser representados por una secuencia infinita de colectivos matemáticos. Este tipo de problema se relaciona con una de las objeciones que más se ha discutido, pues es un problema fundamental en la teoría frecuencialista de von Mises. Según Galavotti, de Finetti en su texto *Filosofía della probabilità* (1979/1995) alega que en la relación entre probabilidad y frecuencionalismo hay dificultades, si bien una de las tareas del cálculo de la probabilidad es demostrar esta relación, considera que von Mises no lo hace, porque toma como supuesto lo que trata de demostrar, por ello, de Finetti afirma que:

> Desde el punto de vista frecuentista, las frecuencias convergen a la probabilidad por definición. Pero desafortunadamente aquí no hay garantía de esto. Los orígenes del frecuentismo desde la mistificación pretenden que las leyes de la probabilidad sean ciertas. [de Finetti en Galavotti 2005, p.221]

El problema que de Finetti intenta exponer es el que de eventos observables se intenta calcular la probabilidad de estos eventos como si fueran secuencias infinitas de eventos. De esta manera, de Finetti apunta a un problema esencial en la teoría de la probabilidad, cómo relacionar los eventos empíricos, finitos, con la característica de ser observados, con el cálculo matemático formal de la teoría de la probabilidad. Independientemente cuál sea su criterio deben de justificar su interpretación y satisfacer los axiomas de la teoría de la probabilidad.

Otra crítica de de Finetti a la teoría frecuencialista de von Mises es su noción de colectivo, pues como vimos en el apartado probabilidad objetiva en el primer capítulo, no es confiable su definición; además, sólo se puede remitir a generalizaciones de casos repetibles, de esta forma, no puede tratar casos singulares particulares que puede haber en la ciencia, casos como hipótesis radicales y que van en contra del pensamiento común.

Otro problema es que sólo podemos elegir una secuencia y limitarnos a sus resultados, es decir, no podemos tener una variedad de secuencias de caso, sino sólo una secuencia en donde realizamos instancias de una misma prueba. Este problema es más importante en el ámbito científico, pues se necesitan hacer varias pruebas a una misma hipótesis para saber si es probable o improbable, de este modo utilizar un método probabilístico que desde sólo una prueba genere instancias que calcularán la probabilidad de una hipótesis, es visto como deficiente.

Una limitación de la teoría frecuencialista es su restricción a un conjunto de eventos uniformes, es decir, sólo trata con colectivos; en cambio, la aproximación subjetiva da un método para medir las probabilidades tomando en cuenta grados de creencia, esta forma de calcular la probabilidad demostró que fue posible extender la probabilidad cuantitativa y que el cálculo matemático puede ser aplicado en casos donde se carece de colectivos. No es que la teoría frecuencialista de von Mises sea incorrecta, sino sólo limitada a eventos cuantitativos. Sin embargo, es posible usar la teoría frecuencialista en los métodos subjetivos (por ejemplo, para calcular la probabilidad *a priori*, como propone Salmon) para extender el cálculo matemático a ejemplos que escapan del campo de la teoría frecuencialista.

Una eficiencia que tiene el método frecuencialista es la limitación de la definición de probabilidad, su definición excluye algún uso del lenguaje ordinario de la probabilidad por el cual un colectivo no puede ser definido y esto, en un sentido, es beneficioso porque excluye algunos usos ordinarios del lenguaje de la probabilidad, los cuales son inadecuados para la interpretación matemática. Von Mises ejemplifica este beneficio:

> La probabilidad de ganar una batalla, por ejemplo, no tiene lugar en nuestra teoría de la probabilidad, porque no podemos pensar en un colectivo en el cual pertenezca. La teoría de la probabilidad no puede ser aplicada a este problema más que el concepto físico de trabajo se puede aplicar al cálculo de 'trabajo' realizado por un agente de recitar su parte en una obra de teatro. [Von Mises en Gillies 2000, p. 98]

Este problema también es visto como el problema de la referencia de clases, pues no tenemos un criterio para saber a qué clase pertenece tal evento. Por ejemplo, supongamos que se quiere saber la probabilidad de que una computadora funcione en los últimos meses de su garantía; para aplicar la teoría frecuencialista se necesita ubicar la clase a donde pertenece el evento; para hacerlo se necesita saber cuál referencia de clases se debe de consultar. La respuesta puede ser cualquiera, es decir, se puede consultar tanto a la clase de las computadoras de ciertas características o a la clase de todas las computadoras; un caso individual puede ser referido a cualquier clase desde varios criterios. Otro ejemplo podría ser, si seguimos nuestro ejemplo de saber cuál es la probabilidad de que un individuo en especifico tenga la herencia genética *y*, en este caso qué debemos considerar es algo que sólo los expertos en este tema pueden decidir y aun así habrá quienes tengan razones para decir, por ejemplo, que considerar en donde vivió, para algunos sería importante para saber si es portador de la herencia genética y, en cambio para otros no es relevante saber en donde vivió sino en donde vive ahora, en qué población.

La teoría frecuencialista es una interpretación objetiva de la probabilidad, pero como vimos con Hanson, hay muchas dificultades para concluir que los eventos pueden ser vistos de manera objetiva. En la investigación científica se sabe de los problemas que implica tratar fenómenos del mundo; un método frecuencialista objetivo tiene la deficiencia de no considerar las diversas interpretaciones de los hechos observados, aunque tiene la eficiencia de ser muy restringido.

El método frecuencialista ha predominado en el campo del análisis de datos estadísticos, este tipo de método se caracteriza por el proceso formal y objetivo que realiza de los datos observados a las conclusiones, un problema que tiene este método es cuando se trata de aplicar para corroborar hipótesis científicas, pues los datos observados es evidencia empírica que puede ser interpretada de manera distinta y puede ser considerada como una interpretación relevante que afecte el cálculo de probabilidad, pues, no puede ser totalmente objetivo. Además, tiene la deficiencia de no poder aplicarse a muestras demasiado pequeñas, pues necesita una secuencia muy grande para calcular la probabilidad.

En conclusión, podemos decir que las ventajas o desventajas que hay en usar el método frecuencialista en el método hipotético deductivo científico en la corroboración de hipótesis científicas, se basan en la forma en cómo calcula la probabilidad de los eventos y, por consiguiente, a la clase de eventos que pueden calcularse sin tanto problema en el método frecuencialista; esta clase de eventos tienen la característica de ser cuantitativos, en cambio, cuando aplicamos el método frecuencialista a eventos cualitativos tenemos varios problemas que ya mencionamos. Por ello, las ventajas o desventajas de usar tal teoría de la probabilidad en un método científico para la corroboración de hipótesis científicas depende mucho en dónde se aplique, a qué tipo de eventos y cómo se aplique.

V. Conclusiones

En este capítulo se confrontarán las razones que puede haber a favor o en contra del uso de cualquiera de los dos métodos de probabilidad ya citados a lo largo de este texto; esto se hará con el fin de sustentar la hipótesis principal de este texto, a saber, el método bayesiano de la probabilidad tiene ventajas sobre el método frecuencialista de la probabilidad en ciertas circunstancias y en lo métodos científicos, específicamente en los métodos científicos que tienen la estructura causa-efecto, en particular el método hipotético deductivo, para corroborar o refutar hipótesis.

A lo largo de este texto hemos visto las distintas teorías de la probabilidad más comunes; para nuestro interés, he destacado dos de ellas: la teoría frecuencialista de la probabilidad y la teoría bayesiana de la probabilidad[105]. En los primeros capítulos he expuesto de una manera formal las teorías, después expuse varias interpretaciones filosóficas sobre la probabilidad, las consecuencias de su uso en la vida común y en la

---

[105] En el texto [Colin Howson y Peter Urbach 2006, p.305] podemos notar que en su conclusión afirma que considera la teoría bayesiana de la probabilidad como un razonamiento probabilístico consistente, la probabilidad la llaman como un razonamiento incierto en donde la fuerza de sus premisas son consideraciones exógenas. Sin embargo consideran que aquí hay un debate sobre la asignación coherente y sostenible de una interpretación de la probabilidad.

investigación científica. Hemos destacado su uso como métodos probabilísticos para complementar el método científico que es usado para corroborar o refutar hipótesis científicas.

Se ha llegado a la conclusión de que el método bayesiano de la probabilidad tiene más ventajas por sus usos que el método frecuencialista de la probabilidad en ciertas circunstancias, específicamente como instrumento para corroborar o refutar hipótesis científicas utilizando el método hipotético deductivo. Una de las razones que tenemos para apoyar la conclusión es que el método bayesiano de la probabilidad puede ser aplicado tanto a eventos cualitativos como a eventos cuantitativos[106], es decir, puede ser utilizado tanto en las ciencias empíricas como en las formales. Además, hemos visto que en los métodos científicos que proceden de la misma forma que el método hipotético deductivo para corroborar hipótesis, las decisiones que se toman sobre la corroboración de hipótesis científicas tienen características cualitativas, pues el manejo de evidencia con características cualitativas y los criterios para elegir qué método usar son interpretaciones que el cálculo de la probabilidad matemática no puede manejar, ya que sólo se restringe al ámbito cuantitativo. También, cuando pensamos en la probabilidad como una herramienta que es usada en los métodos científicos, debemos pensar los problemas que la probabilidad enfrenta como problemas en parte a datos empíricos (cualitativos) y en parte a datos aritméticos (cuantitativos). La probabilidad enfrenta este tipo de problemas; un problema que se deriva del uso de la probabilidad es que no sabemos hasta qué punto es asunto de la probabilidad tratar con asuntos aritméticos ya sea como una herramienta que se usa en los

[106] Hacking menciona una distinción entre las probabilidades, esta distinción considera una probabilidad como cualitativa, si trata eventos no numerables, y considera una probabilidad cuantitativa, aquella que trata sólo con eventos numerables.

métodos científicos para decidir si una hipótesis debe ser corroborada o como un método probabilístico que maneja datos empíricos y no empíricos para calcular la probabilidad de un evento que puede contener características cualitativas. Hemos de decir que la probabilidad comenzó como un método para tratar con problemas azarosos, con problemas que tenían características cualitativas.

Este tipo de problemas en el que no sabemos qué le corresponde a la probabilidad analizar, a saber, sólo datos aritméticos o también datos empíricos. Esta incertidumbre afecta en cómo debemos usar el método de Bayes, pues puede ser interpretado bajo varias posturas de la probabilidad como las que hemos explicado: la subjetiva y la objetiva. Aunque podría pensarse que el método bayesiano presupone la interpretación subjetiva, ya que fue desarrollado por Thomas Bayes, quien asumió la interpretación subjetiva de la probabilidad, esto no es del todo correcto: como hemos mostrado, se puede usar el método bayesiano tanto bajo una interpretación subjetiva de la probabilidad como bajo una interpretación objetiva de la probabilidad. Por ejemplo, podemos usar el método bayesiano para calcular la probabilidad de eventos con características cualitativas o eventos que suceden muy pocas veces por varias circunstancias (económicas o imposibles de repetir), en este tipo de eventos en donde no se cuenta con una secuencia considerable para calcular la probabilidad, cuando usemos el método frecuencialista, pues tenemos una secuencia, también se puede recurrir también al método de Bayes. Por otra parte, el método bayesiano puede usar el método frecuencialista para calcular la probabilidad *a priori.*

En contraste, el método frecuencialista de la probabilidad, sólo se puede usar bajo una interpretación objetiva de la probabilidad, pues las definiciones que asume deben ser satisfechas para utilizar el método frecuencialista de una manera correcta, es decir, es necesario asumir la interpretación objetiva de la probabilidad. Por consiguiente, sólo podemos aplicar el método frecuencialista si primero tenemos una experiencia de éxito o fracaso sobre la hipótesis en cuestión, entre otras determinaciones.

Si para aplicar el teorema de Bayes se adopta una interpretación objetiva de la probabilidad, y se aplica el método frecuencialista, no podremos aplicarlo hasta tener alguna experiencia con el éxito o fallo de la hipótesis en general, aunque el método frecuencialista es más formal, riguroso y restringe las varias posibles hipótesis, pierde utilidad en el ámbito cualitativo, pues hay hipótesis científicas para las que es muy riesgoso o muy costoso hacer experimentos y de esta manera poder determinar la probabilidad de una hipótesis; recordemos el caso del experimento de la colisión entre dos átomos.

Sin embargo, el método frecuencialista, comparado con la teoría bayesiana subjetiva, tiene la ventaja de ser un método completamente formal, objetivo y riguroso, lo que es una garantía de que evita ambigüedades como las que se originan en la interpretación subjetiva de la probabilidad del método bayesiano.

No obstante, el método de Bayes puede ser utilizado como un método completamente riguroso y formal en la probabilidad cualitativa como en la cuantitativa, porque como vimos en el texto, se expuso una manera en el que el elemento subjetivo en la aplicación del teorema de Bayes ha sido circunscrito en la probabilidad cuantitativa, en ambas probabilidades, cuantitativa y cualitativa se asume el método bayesiano. En la corroboración o refutación de las hipótesis científicas por medio de la observación de la evidencia es útil usar un método que se adapte a la evidencia con características cualitativas y a la dificultad que se relaciona con la observación de la evidencia.

Además, se ha dicho que el método bayesiano complementa, de una manera eficaz, el método hipotético deductivo científico, porque la estructura causa-efecto que tiene es muy semejante a la forma del método científico en general, como Salmon afirma: "el análisis de las inferencias por las cuales las hipótesis científicas son corroboradas por evidencia observacional, muestra, yo creo, que esta estructura es dada por el teorema de Bayes"[107]. Luego, la forma en que se confirma una hipótesis científica que parte de los datos observados que ya sucedieron y afirmar las hipótesis, pueden ser comparadas con el efecto de asumir tales hipótesis, y de ahí tratamos de explicar qué pudo haber causado tal evento. Esta forma de corroborar hipótesis por lo general es la manera en que los métodos científicos proceden; por ello, podemos concluir que esta es otra razón por la que el método de Bayes tiene más utilidad en los métodos científicos con la estructura de causa-efecto.

Otra característica del método bayesiano es que su método para calcular probabilidades utiliza varias probabilidades, es decir, calcula la probabilidad total (que vimos en el primer capítulo), la de la evidencia y la probabilidad de la evidencia dada la hipótesis, para hacer estos cálculos utiliza probabilidad condicional, del producto y probabilidad total[108]. En cambio el método frecuencialista por ser más sencillo su método sólo calcula la probabilidad de la hipótesis dada la evidencia; por medio de instancias calcula la probabilidad de la hipótesis.

---

[107] Este es el texto original que traduje: The analysis of the inference by which scientific hypotheses are confirmed by observational evidence shows, I believe, that its structure is given by Bayes' theorem. [Salmon 1984, p.131]

[108] En el texto de [Colin Howson y Peter Urbach 2006, pp.13-114] se puede consultar una explicación formal de las varias funciones de probabilidad que usa el método bayesiano, también se puede consultar la explicación formal del método frecuencialista de la probabilidad.

Hemos visto que la teoría de Bayes también tiene igual forma de calcular la probabilidad que los métodos científicos con la estructura de causa-efecto, ya que comienza por cuestionarse cuál es la probabilidad de que sea tal hipótesis la causa de tal evento dada la probabilidad de los datos que ya sucedieron, es decir, cuál es la probabilidad de una hipótesis *A* dado *B*, donde *B* es el evento.

De esta manera, en los métodos científicos a partir de los datos se trata de corroborar una hipótesis que los explique, así es como calcula la probabilidad el método bayesiano, a partir de la probabilidad de los datos se intenta calcular la probabilidad de las hipótesis alternativas.

Otra ventaja del método bayesiano es que se utiliza, por lo general, en varios campos de investigación científica: inteligencia artificial, astronomía, medicina, física, geofísica, entre otras ciencias que usen probabilidad y un método científico con la estructura de causalidad[109].

Un problema que tiene el método frecuencialista y que se le objeta mucho es la relación que establece entre observación y frecuencia, pues como vimos, define frecuencia de una forma, pero no define observación de evidencia, pues tratar con la observación de la evidencia es relacionar el ámbito cualitativo y evaluar epistemológicamente. Además, como hemos visto, en ciertas circunstancias, cuando se usa el método frecuencialista no se puede

[109] Sin embargo, he de reconocer que el método frecuencialista tiene una mejor aplicación en las ciencias que usan más estadísticas, pues la estadística trata con eventos que satisfacen las definiciones que la teoría frecuencialista que esta teoría exige para ser usada y obtener resultados objetivos, demás usan eventos que tienen una frecuencia y pertenecen a una clase que son requisitos indispensables para usar esta teoría.

realizar varias pruebas a una hipótesis científica para calcular su probabilidad, pues sólo se limita a calcular la probabilidad de instancias de una sola prueba.

Por último, otra línea de investigación que veo muy apegada al problema de la utilidad de los métodos de probabilidad, específicamente bayesiano y frecuencialista, es la postura de Salmon, ya que debemos decir que la probabilidad se usa para descartar las hipótesis plausibles pero menos probables, esto lo hace para restringir más las hipótesis que deben ser consideradas. Independientemente de qué método probabilístico se considere que tenga más aplicaciones, aunque ambos métodos probabilísticos tienen ventajas y desventajas, pero, coincido con Salmon en que tendríamos un método probabilístico más consistente, sencillo, productivo y que tuviera más utilidad si se complementan entre ellos, pues el método bayesiano, aunque tenga la ventaja de que su probabilidad *a priori* no es fija, tiene un problema fundamental, el cual es la estimación de dicha probabilidad en este caso puede usarse el método frecuencialista para calcularla, de esta forma podemos obtener un nuevo y más útil método. La idea es tratar de construir un método probabilístico que sea útil para el método científico, tomar las ventajas que tienen los métodos probabilísticos y tratar de combinarlas, crear un método probabilístico que pueda ser usado tanto en eventos cualitativos como en eventos cuantitativos.

Asignar probabilidades es una forma de evaluar epistemológicamente para saber qué teoría científica rechazar y cuál corroborar, por ello, depende de cómo se aplique y a qué tipo de eventos y con qué método probabilístico, ya sea bayesiano o frecuencialista. En los métodos científicos debemos ser conscientes de la dificultad que involucra la observación de la evidencia, pues como se interprete se darán datos por medio de los cuales se calcula la probabilidad de la hipótesis científica que se quiera calcular, de ello depende además de la experiencia del científico que es experto en el área en donde corresponde la hipótesis que se quiere calcular la probabilidad de ello depende que resulte acorde a las creencias o no. La ciencia trata de explicar los fenómenos del mundo, y las pruebas que se hagan para la corroboración de una hipótesis, como hemos visto, son tanto cualitativas como cuantitativas; éstas deben ser consideradas por el método probabilístico que se quiera aplicar.

Bibliografía

Aliseda Atocha and Gillies Donald. 2007. "Logical, Historical and Computational Approaches." en Kuipers. T. *General Philosophy of Science: Focal Issues.* (Handbook of the Philosophy of Science, Vol. I). Amsterdam: North Holland. Elsevier.

Bovens, Luc. 2003. *Bayesian epistemology.* Oxford: Clarendon.

Cohen, Jonathan L. 1989. *The Philosophy of Induction and Probability.* New York: Oxford University Press.

Collin Howson y Peter Urbach. 2006 *Scientific reasoning: the Bayesian approach.* Estates United American: Chicago and La Salle, Illinois Open Court.

Dale Andrew. I. 1991. *A History of Inverse Probability from Thomas Bayes to Karl Pearson*. New York: Spinger-Verlag.

Diez, A. y Moulines, U. 1999. *Fundamentos de Filosofía de la Ciencia*. Barcelona: Ariel.

Galavotti, María C. 2005. *Philosophical introduction to probability.* Stanford, California: CSLI.

García Álvarez, M. A. 2005. *Introducción a la teoría de la probabilidad.* Primer curso. México: FCE.

__________________2005. *Introducción a la teoría de la probabilidad.* Segundo curso. México: FCE.

Gillies, Donald. 2000. *Philosophical theories of probability.* London and New York: Routledge.

Hacking Ian. 1995. *El surgimiento de la probabilidad.* Barcelona: Gedisa..

__________ 2001. *An Introducción to Probability and Inductive Logic*. New York: Cambridge University Press.

Hanson. N. R. 2005. "Observación" en *Filosofía de la ciencia: teoría y observación*. Comp. León Olivé y Ana Rosa Pérez Ransanz. 2da Ed. México: Siglo XXI/ UNAM.

Hernández Ramírez, M. A. 2007. *La noción de posibilidad en tres definiciones matemáticas de probabilidad.* México: UNAM.

Howie, David. 2002. *Interpreting probability: controversies and developments in the early twentieth century*. Cambridge: Cambridge University Press.

Hume, David. 2004. *Investigación sobre el entendimiento humano.* (traducido por: Vicente Sanfélix Vidarte y Carmen Orqs Marqués). Madrid: Istmo.

Karlin Samuel. 1975. *A first course in stochastic processes*. 2da ed. New York: Stanford University Academic Press.

Keynes, Jhon Maynard. 2004. *A treatise on probability*. Mineloa, New York: Dover.

Kneale, William. 1949. *Probability and Induction*. Oxford: Oxford University Press.

Kolmogorov, A.N. 1956. *Foundations of Probability and physics*. New York: Chelsea publishers.

Kuipers, Theo. 2007. *General philosophy of science: focal issues*. Amsterdam: North Holland.

Laplace Pierre, Simon. 1994. *Philosophical essay on probabilities.* Nueva York (Traducida por Dale Andrew de la versión francesa de 1825) En Dale 1991. *A History of Inverse Probability from Thomas Bayes to Karl Pearson*. New York: Spinger-Verlag.

Maistrov, L.E. 1974. *Probability Theory.* Nueva York (Traducida por S. Kotz de la versión rusa de 1964) En Dale 1991. *A History of Inverse Probability from Thomas Bayes to Karl Pearson*. New York: Spinger-Verlag.

Popper, Karl. 1977. *La lógica de la investigación científica*. Madrid: Tecnos.

Russell Bertrand. 1926. *Our knowledge of the external world.* London: George Allen & Unwin.

_____________ 1992. *Human Knowledge: its scope and limits*. London: Routledge.

Salmon Wesley. 1984. *The Foundations of Scientific Inference.* Pittsburg: University of Pittsburg Press.

____________ 1973. *Logic.* Englewood Cliffs: Prentice Hall.

Venn, Jhon. 2006. *The Logic of chance.* Mineloa, NY: Dover.

Von Mises, Richard. 1981. *Probability, Statistics and Truth.* (editada y compilada por Hilda Geiringer). New York: Academic.

_______________ 1964. *Mathematical theory of probability and statistics.* (editada y Compilada por Hilda Geiringer). New York: Academic.

Von Plato, Jan. 1994. *Creating Modern Probability*: Cambridge: Cambridge University Press.