Diseño por Eliminación Versus Diseño por Comparación

(Capítulo 33 de *The Design Revolution*)

William A. Dembski

¿Cómo se infieren propiamente las hipótesis de diseño? ¿Simplemente eliminando hipótesis de azar o comparando la verosimilitud del azar y las hipótesis de diseño?

DETRÁS DE ESTA PREGUNTA HAY DOS APROXIMACIONES fundamentalmente diferentes acerca de cómo razonar con las hipótesis de azar, una amigable con el diseño inteligente y la otra no tanto. La aproximación amigable, debida a Ronald Fisher, rechaza una hipótesis de azar dado que los datos muestrales aparezcan en una región de rechazo pre-especificada. La aproximación poco amigable, debida a Thomas Bayes, rechaza una hipótesis de azar dado que una hipótesis alternativa confiere una mayor probabilidad a los datos en cuestión cuando comparada con la hipótesis original. En la aproximación fisheriana, las hipótesis de azar se rechazan de manera aislada cuando bajo esas hipótesis los datos son muy improbables. En la aproximación bayesiana las hipótesis de azar se eliminan dado que otras hipótesis hacen más probable al conjunto de datos. Mientras que en la aproximación fisheriana el énfasis está en la eliminación, en la aproximación bayesiana el énfasis está puesto en la comparación. Estas dos aproximaciones son incompatibles y la misma comunidad estadística está sumida en una discusión sobre cuál de estas aproximaciones adoptar como canon correcto de racionalidad estadística. La diferencia refleja una divergencia profunda en las intuiciones fundamentales sobre la naturaleza de la racionalidad estadística y, en particular, sobre lo que cuenta como evidencia estadística.

La crítica más influyente de la complejidad especificada la acusa con situarse en el lado errado de esta división. Específicamente, los críticos acusan que usar la complejidad especificada para inferir diseño presupone la aproximación fisheriana, eliminativa, para razonar con las hipótesis de azar mientras que la aproximación correcta para inferir diseño debe abrazar la aproximación bayesiana, comparativa. El más prominente académico que hace esta crítica es Elliot Sober. Otros académicos han erigido esta crítica también, y muchos más incluso la han citado como decisiva para refutar que la complejidad especificada sea una señal de inteligencia.

En respuesta a esta crítica permítaseme iniciar con un examen de la realidad. A menudo cuando la literatura bayesiana intenta justificar sus métodos contra los métodos de Fisher, los autores aceptan rápidamente que los métodos fisherianos dominan el mundo científico. Por ejemplo, Richard Royall (quien, estrictamente hablando es un teórico de verosimilitud en vez de un bayesiano, aunque la discusión no es central para esta discusión) escribe, “las pruebas de hipótesis estadísticas, como son más comúnmente usadas al analizar y reportar resultados de estudios científicos, no proceden... con el hacer una escogencia entre dos [o más] hipótesis especificadas... [sino que siguen] un procedimiento más común” (*Statistical Evidence: A Likelihood Paradigm* [*Evidencia Estadística: Un paradigma de Verosimilitud*], Chapman & Hall, 1997). Seguido de esto, Royall pasa a esbozar ese procedimiento común, el cual requiere especificar una única hipótesis de azar, usar una estadística de prueba para identificar una región de rechazo, chequear si la probabilidad de esa región de rechazo bajo la hipótesis de azar cae debajo de un nivel de significancia dado, determinar si una muestra (los datos) cae dentro de esa región de rechazo y, si es así, rechazar la hipótesis de azar. En otras palabras, las ciencias miran a Fisher y no a Bayes para su metodología estadística. Colin Howson y Peter Urbach, en *Scientific Reasoning: The Bayesian Approach* [*Razonamiento Científico: La Aproximación Bayesiana*], admiten de manera semejante la abrumadora falta de popularidad de los métodos bayesianos entre los científicos cuando están trabajando.

¿Entonces está siendo la mayoría de los científicos simplemente estúpida o perezosa al adoptar la aproximación fisheriana para el razonamiento estadístico? Para responder esta pregunta vamos a mirar dos ejemplos prototípicos en los cuales se emplean los métodos fisheriano y bayesiano. Una vez tengamos estos ejemplos a la mano, podemos utilizarlos para ver lo que puede salir mal con ambos métodos. Empecemos con un ejemplo del razonamiento fisheriano. El razonamiento fisheriano elimina las hipótesis de azar de manera aislada, así que sólo necesitamos considerar una única hipótesis de azar a eliminar. Tomemos una particularmente simple, a saber, la hipótesis de azar que caracteriza el lanzamiento de una moneda honesta. Para ver si la moneda está sesgada a favor de las caras (luego no es honesta), se puede determinar una región de rechazo de diez caras en serie y luego lanzar la moneda diez veces. En la aproximación de Fisher si la moneda cae cara las diez veces, entonces está justificado el rechazar la hipótesis de azar. La improbabilidad de obtener diez caras en serie, asumiendo que la moneda es honesta, es aproximadamente una en mil (es decir, 0.001).

A continuación, para ilustrar la aproximación bayesiana, considere el siguiente esquema probabilístico. Imagine dos monedas, una honesta y la otra sesgada. Suponga que la moneda sesgada tiene probabilidad de caer cara el 90% de las veces. Adicionalmente, imagine una urna gigante con un millón de bolas de igual tamaño, donde todas son blancas excepto una que es negra. Ahora imagine que una única muestra aleatoria se tomará de la urna y si una bola blanca es seleccionada (lo cual es abrumadoramente probable) entonces la moneda honesta se lanzará diez veces; pero si se selecciona la bola negra (lo cual es abrumadoramente improbable), entonces la moneda sesgada se lanzará diez veces. Ahora imagine que todo lo que usted ve es que la moneda se lanzó diez veces y todas las veces cayó cara. La probabilidad de que los lanzamientos resultaran en diez caras en serie, dado que la moneda es honesta, es aproximadamente .001 (una en mil). Pero la probabilidad de que caigan diez caras en serie, dado que la moneda sesgada se lanzó, es aproximadamente .35 (un poco mejor que una en tres). Dentro de la literatura bayesiana estas probabilidades se conocen como *verosimilitudes*.

Entonces ¿cuál moneda se lanzó, la honesta o la sesgada?. Si se miran solamente las verosimilitudes, parecería que la moneda sesgada fue la lanzada; en realidad, es mucho más verosímil que aparezcan diez caras en serie si se usa la moneda sesgada en vez de la moneda honesta. Pero esa respuesta no serviría. El problema es que cuál moneda es lanzada tiene lo que en la literatura bayesiana se conoce como *probabilidad a priori*. La probabilidad a priori hace mucho más probable que la moneda honesta fuera lanzada y no que lo fuera la moneda sesgada. La moneda honesta tiene probabilidad a priori .999999 de ser lanzada (porque una bola blanca tiene esa probabilidad de ser seleccionada en la urna), mientras que la moneda sesgada tiene probabilidad a priori .000001 de ser lanzada (porque la única bola negra tiene esa probabilidad de ser seleccionada en la urna).

Para decidir cuál de las monedas fue lanzada, estas probabilidades a priori deben descomponerse en factores dentro de las verosimilitudes calculadas anteriormente. Para hacer esto, se calcula lo que en la literatura bayesiana se conoce como *probabilidad a posteriori* (la cual se obtiene utilizando el teorema de Bayes). La probabilidad a posteriori de que la moneda honesta se haya lanzado dado que se observaron diez caras en serie es .9996, mientras que la probabilidad a posteriori de que la moneda sesgada haya sido lanzada dado que se observaron diez caras en serie es .0004. Por lo tanto, dado el esquema probabilístico de las dos monedas y la urna como se describió anteriormente, es mucho más probable que la moneda honesta hubiera sido lanzada y no la moneda sesgada. Y este es el caso aun cuando el resultado observado de las diez caras en serie sea en sí mismo más consistente con la moneda sesgada que con la moneda honesta.

Dadas estas ilustraciones particularmente nítidas y claras de las aproximaciones fisheriana y bayesiana, se podría preguntar cuál es el problema con cada una. Ambas aproximaciones, como ilustradas en estos ejemplos, parecen eminentemente razonables dadas las preguntas que están llamadas a contestar. No obstante, de ambas aproximaciones surgen serios problemas conceptuales cuando se escudriñan más a fondo. Quiero, en lo que queda de este capítulo, describir los problemas conceptuales que surgen de la aproximación fisheriana e indicar cómo mi trabajo en complejidad especificada ayuda a resolverlos. Lo siguiente que quiero hacer es describir los problemas conceptuales que surgen de la aproximación bayesiana e indicar porqué son inadecuados como modelo general para el razonamiento estadístico. En particular, mostraré que la aproximación fisheriana puede hacerse lógicamente coherente y porqué la aproximación bayesiana, cuando funciona (lo cual no sucede muy a menudo), debe presuponer, en efecto, la aproximación fisheriana.

Entonces ¿cuáles son los problemas con la aproximación fisheriana y cómo ayuda a resolverlos mi trabajo en complejidad especificada? Esquemáticamente la aproximación fisheriana luce así: una hipótesis de azar definida con respecto a una clase de referencia de posibilidades es dada. También es dada una región de rechazo en esa clase de referencia. Con la hipótesis de azar y la región de rechazo a la mano, se procede a muestrear un evento de la clase de referencia de posibilidades. Si ese evento (la muestra o los datos) cae dentro de una región de rechazo, y si la probabilidad de esa región de rechazo con respecto a la hipótesis de azar es suficientemente pequeña, entonces se rechaza la hipótesis de azar. Intuitivamente, piense en el lanzamiento de flecha a un muro grande que tiene un blanco fijo. El muro corresponde a la clase de referencia de posibilidades (todos los lugares a los cuales la flecha puede llegar) y el blanco corresponde a la región de rechazo. Si que la flecha aterrice en el blanco (es decir, la muestra cae en la región de rechazo) tiene probabilidad lo suficientemente pequeña, entonces se rechaza la hipótesis de azar. En nuestro ejemplo anterior de los lanzamiento de moneda, la clase de referencia era todas las posibles sucesiones de caras y sellos, la región de rechazo era todas las sucesiones con diez caras en serie, la muestra era la sucesión de diez caras en serie y la hipótesis de azar presuponía que la moneda era honesta.

¿Hay algo errado con este cuadro? Aunque el cuadro ha probado ser exitoso en la práctica, cuando Fisher formuló los apuntalamientos teóricos dejó de lado una cosa deseable. Hay tres preocupaciones principales: primero ¿cómo hacer preciso lo que significa para una región de rechazo tener probabilidad “suficientemente pequeña” con respecto a la hipótesis de azar? Segundo ¿cómo se caracterizan las regiones de rechazo de manera que una hipótesis de azar no se rechace automáticamente en caso de que esa hipótesis en realidad esté operando? Y tercero ¿por qué una muestra que cae en una región de rechazo cuenta como evidencia contra una hipótesis de azar?

El primer punto se plantea usualmente en términos de establecer un “nivel de significancia”. Un nivel de significancia prescribe el grado de improbabilidad por debajo del cual una región de rechazo elimina una hipótesis de azar una vez que la muestra ha caído dentro de esa región. Los niveles de significancia en la literatura de ciencias sociales, por ejemplo, usualmente se toman en los valores 0.05 y 0.01. Pero ¿de dónde vienen estos números? A la verdad, son completamente arbitrarios. Esta arbitrariedad ha perseguido la aproximación fisheriana desde su inicio. No obstante, hay una forma de salirle al paso.

Considere de nuevo nuestro ejemplo de lanzar una moneda diez veces y obtener en ese intento diez caras en serie. La región de rechazo que se ajusta a esta sucesión de lanzamientos de moneda establece, por lo tanto, un nivel de significancia de 0.001. Si obtenemos diez caras en serie podemos considerar esto, por lo tanto, como evidencia en contra de que la moneda es honesta. ¿Pero qué sucede si no lanzamos la moneda diez veces no sólo en una ocasión sino que la lanzamos diez veces en repetidas ocasiones? Si el comportamiento de la moneda fuera enteramente lo esperado de una moneda honesta la mayoría de veces que la lanzamos, entonces en las pocas ocasiones en las cuales observemos diez caras en serie, no tendríamos razón para sospechar que la moneda estaba sesgada puesto que las monedas honestas, cuando son lanzadas lo suficientemente a menudo, producirán una sucesión cualquiera de lanzamientos de moneda, incluyendo diez caras en serie. La fuerza de la evidencia contra una hipótesis de azar cuando una muestra cae dentro de la región de rechazo por lo tanto depende de cuántas muestras son tomadas o pueden haber sido tomadas. Estas muestras constituyen lo que llamo *recursos replicacionales*. Entre más muestras de este estilo, mayores los recursos replicacionales.

Los niveles de significancia por lo tanto necesitan tener en cuenta los recursos replicacionales si es que las muestras que alcanzan estos niveles van a tenerse en cuenta como evidencia contra una hipótesis de azar. Pero eso no es suficiente. Además de tener en cuenta los recursos replicacionales, los niveles de significancia también necesitan tener en cuenta lo que llamo *recursos especificacionales*. La región de rechazo en la cual nos hemos estado enfocando especificó diez caras en serie. Pero con certeza si las muestras que caen dentro de esta región de rechazo podrían contar como evidencia en contra de que la moneda es honesta, entonces las muestras que caen dentro de otras regiones de rechazo, de manera semejante, deben contar como evidencia en contra de que la moneda es honesta. Por ejemplo, considere la región de rechazo que especifica diez sellos en serie. Por simetría, las muestras que caen dentro de esta región de rechazo deben contar como evidencia en contra de que la moneda es honesta, tanto como las muestras que caen dentro de la región de rechazo que especifica diez caras en serie.

Pero si este es el caso ¿qué prevendría, entonces, al rango entero de posibles lanzamientos de moneda de ser absorbido por regiones de rechazo de manera que sin importar cuál sucesión de monedas se observó, siempre termine esta cayendo en alguna región de rechazo y contando por lo tanto como evidencia en contra de que la moneda sea legal? Más generalmente ¿qué previene que una clase de referencia de posibilidades cualquiera sea particionada en una colección mutuamente exclusiva y exhaustiva de regiones de rechazo de manera tal que una muestra cualquiera siempre caiga en una de estas regiones de rechazo y por lo tanto cuente como evidencia en contra de una hipótesis de azar cualquiera?

La forma de salirle al paso a este punto es limitar las regiones de rechazo a aquellas que pueden caracterizarse por patrones de baja complejidad (de hecho tal limitación ha estado implícita cuando se aplican en la práctica los métodos fisherianos). Las regiones de rechazo, y las especificaciones más generalmente, corresponden a eventos y por lo tanto tienen asociada una probabilidad o complejidad probabilística. Pero las regiones de rechazo también son patrones, y como tales tienen una complejidad asociada que mide el grado de complicación de los patrones, o lo que yo llamo *complejidad especificacional*. Usualmente esta forma de complejidad corresponde a la medida de compresibilidad de Kolmogorov o la longitud de descripción mínima (entre más corta sea la descripción, más baja es la complejidad especificacional. Vea <www.mdl-research.org>). Ya resumí estos dos tipos de complejidad en el capítulo 10. Note que la complejidad especificacional surge muy naturalmente: no es una construcción ad hoc o artificial diseñada simplemente para salvar la aproximación de Fisher. Más bien, ha estado implícita desde hace tiempo, permitiéndole a la aproximación de Fisher florecer a pesar de los apuntalamientos teóricos inadecuados con los cuales Fisher la dotó.

Los recursos replicacionales y especificacionales juntos constituyen lo que llamo *recursos probabilísticos*. Los recursos probabilísticos resuelven las primeras dos preocupaciones (mencionadas anteriormente) concernientes a la aproximación de Fisher para el razonamiento estadístico. Específicamente, los recursos probabilísticos nos permiten establecer niveles de significancia racionalmente justificados y restringen el número de especificaciones, previniendo con esto que las hipótesis de azar se eliminen de buenas a primeras. Los recursos probabilísticos proveen por lo tanto un fundamento racional para la aproximación fisheriana de razonamiento estadístico. Lo que es más, al estimar los recursos probabilísticos disponibles en el universo físico conocido, podemos establecer un nivel de significancia que está justificado sin importar los recursos probabilísticos en cualquier circunstancia dada. Tal nivel de significancia, independiente del contexto, es de este modo aplicable universalmente y responde definitivamente a qué significa que un nivel de significancia sea “lo suficientemente pequeño” sin importar la circunstancia. Para una estimación conservadora de este nivel de significancia, conocida como una *cota de probabilidad universal*, vea el capítulo diez. Para los detalles acerca de cómo colocar la aproximación fisheriana de razonamiento estadístico sobre un fundamento racional firme, vea el capítulo dos de mi libro *No Free Lunch*.

Eso deja la tercera preocupación concerniente a la aproximación fisheriana de razonamiento estadístico: ¿por qué una muestra que cae en una región de rechazo (o, más generalmente, un resultado que se ajusta a una especificación) debe contar como evidencia en contra de una hipótesis de azar? Una vez se admite que la aproximación fisheriana es lógicamente coherente y que se pueden eliminar las hipótesis de azar individualmente solo con chequear si las muestras caen dentro de regiones de rechazo adecuadas (o, más generalmente, si un resultado se ajusta a especificaciones adecuadas), es asunto sencillo extender este razonamiento a familias enteras de hipótesis de azar, llevar a cabo una inducción eliminativa (ver el capítulo treinta y uno) y eliminar con ello todas las hipótesis de azar relevantes que puedan explicar una muestra. Y de ahí no hay sino un pequeño paso para inferir diseño.

Vamos a quedarnos en el último punto por un momento: ¿cómo se pasa de eliminar el azar para inferir diseño? De hecho ¿qué justifica este paso de eliminación del azar a inferencia de diseño? Estamos suponiendo, por el momento, que la aproximación fisheriana puede eliminar legítimamente hipótesis de azar individuales y de este modo, por eliminación sucesiva, eliminar familias completas de hipótesis de azar. Para eliminar una hipótesis de azar, la aproximación fisheriana determina si un resultado se ajusta a una especificación y si la especificación misma describe un evento de probabilidad pequeña (el evento aquí comprende todos los resultados que se ajustan a la especificación). Dado que hayamos caracterizado con éxito todas las hipótesis de azar que excluyen diseño y que hayamos podido eliminarlas por medio de tal especificación (el resultado exhibe por lo tanto complejidad especificada) ¿por qué deberíamos pensar que el resultado es diseñado?

En este caso la especificación misma actúa como un puente lógico entre la eliminación del azar y la inferencia de diseño. Aquí esta el raciocinio: si podemos señalar un patrón dado de manera independiente (es decir, una especificación) en algún resultado observado, y si los posibles resultados que se ajustan a ese patrón son, tomados conjuntamente, altamente improbables (en otras palabras, si el resultado observado exhibe complejidad especificada), entonces es más plausible que algún agente o proceso que se dirigiera a un fin produjera el resultado al conformarlo intencionalmente al patrón, en lugar de que el resultado terminara ajustándose al patrón simplemente por azar. De acuerdo con esto, aun cuando la complejidad especificada establece diseño por medio de un argumento eliminativo, no es justo decir que esta establece diseño *puramente* por medio de un argumento eliminativo. El patrón dado de manera independiente, o la especificación, contribuye positivamente a nuestro entendimiento del diseño inherente en las cosas que exhiben complejidad especificada.

Para evitar esta resbalosa pendiente al diseño, los teóricos bayesianos niegan que la aproximación fisheriana pueda eliminar legítimamente aunque sea una hipótesis de azar (mucho menos barrer con todas las hipótesis de azar relevantes, como se requiere para una exitosa inferencia de diseño). El problema, como ellos lo ven, es que las muestras que caen en las regiones de rechazo (o, más generalmente, los resultados que se ajustan a una especificación) no pueden servir como evidencia en contra de las hipótesis de azar. Más bien, la única forma de que haya evidencia contra una hipótesis de azar es que haya mejor evidencia a favor de otras hipótesis.

Voy a analizar la aproximación bayesiana para la evidencia estadística momentáneamente, pero primero necesito decir algo acerca de la evidencia en general. En *World Without Design* [*El Mundo sin Diseño*], Michael Rea anota, “la verdadera indagación es un proceso en el cual intentamos revisar nuestras creencias sobre la base de lo que consideramos evidencia”. Continúa él:

> Pero esto significa que, para indagar sobre algo, debemos ya estar dispuestos a tomar algunas cosas como evidencia. Incluso para poder empezar a indagar, debemos tener ya varias disposiciones para confiar al menos en algunas de nuestras facultades cognitivas como fuentes de evidencia y tomar ciertas clases de experiencias y argumentos como evidencia. Tales disposiciones (vamos a llamarlas *disposiciones metodológicas*) pueden ser adquiridas reflexiva y deliberadamente.

De acuerdo con esto, lo que cuenta como evidencia (y eso incluye la evidencia estadística) se decide no sobre la base de la evidencia sino sobre la base de las disposiciones tales que ellas mismas no son gobernadas por la evidencia. ¿Por qué, por ejemplo, la mayoría de los matemáticos encuentran que las pruebas por contradicción (esto es, *reductio ad absurdum*) son evidencia que cuenta para la verdad de una proposición matemática pero otros (los intuicionistas) encuentran que tales pruebas son inadecuadas y requieren en su lugar pruebas constructivas? O de nuevo, ¿por qué las aproximaciones de Fisher y Bayes a la evidencia estadística se mantienen en disputa? En tales casos el debate no es solamente acerca de cómo pesar cierta evidencia sino sobre lo que en principio cuenta como evidencia. El asunto de lo que cuenta como evidencia corta a través de todo el debate sobre diseño inteligente. ¿Puede si quiera existir tal cosa como evidencia a favor de una inteligencia no evolucionada que diseñe la complejidad biológica? Muchos científicos y filósofos naturalistas niegan que pueda existir. Pero para negarlo coherentemente se necesita un marco evidencial. El marco preeminente en este asunto es el bayesiano. Quiero por lo tanto examinar ese marco a continuación y, específicamente, mostrar porqué es tan inadecuado para sacar inferencias de diseño como para excluirlas.

Cuando la aproximación bayesiana intenta juzgar entre las hipótesis de azar y diseño, esta trata a las dos hipótesis como si tuvieran probabilidades a priori y confiriendo probabilidades a los resultados y eventos. De este modo, dada la hipótesis de azar H, la hipótesis de diseño D y el resultado E, el teórico bayesiano intenta comparar las probabilidades a posteriori de H y D con respecto a E (esto es, $P(H \mid E)$ contra $P(D \mid E)$). Si la probabilidad a posteriori de D sobre E es mayor que la de H sobre E, entonces E sirve de evidencia a favor de D y la fuerza de la evidencia es proporcional a cuán grande es $P(D \mid E)$ con respecto a $P(H \mid E)$. Desafortunadamente, calcular las probabilidades a posteriori requiere conocer las probabilidades a priori (es decir, $P(H)$ y $P(D)$), y a menudo esas no están disponibles. En ese caso, se pueden calcular solamente la verosimilitud de E con respecto a H y D (es decir, $P(E \mid H)$ contra $P(E \mid D)$).

Existe una versión minimal de la aproximación bayesiana conocida como la *aproximación de verosimilitud* que esencialmente ignora las probabilidades a priori y nada más mira la razón de verosimilitud (esto es, $P(E \mid H)/P(E \mid D)$) para determinar la fuerza de la evidencia a favor de una hipótesis. Sin embargo, esto va en pro de un entendimiento idiosincrático de la evidencia. La evidencia, como se entiende usualmente, se refiere a lo que nos causa revisar nuestras creencias. Pero las razones de verosimilitud no están en posición de hacer eso sin la ayuda de las probabilidades a priori. Por ejemplo, si yo oigo en mi ático el repiqueteo de pequeños pies y el sonido de pines de bolos cayendo, la verosimilitud de la hipótesis de diseño de que gremlins estén jugando bolos en mi ático puede ser mayor que la verosimilitud de cualquier hipótesis de azar que pretenda explicar esos sonidos. Con todo, mi incredulidad en la hipótesis de los gremlins permanecería tan absoluta y completa como antes por causa de mi creencia previa de que los gremlins no existen (en términos bayesianos, la probabilidad a priori P(D), donde D es la hipótesis del gremlin, para mi completamente cero).

Acabo de describir la aproximación bayesiana que establece la evidencia a favor de las hipótesis de diseño en comparación con las hipótesis de azar. De acuerdo con esto, hacer una inferencia de diseño es determinar que la evidencia, construida en términos bayesianos o de verosimilitud, favorece al diseño sobre el azar. ¿Qué hay de malo con esta aproximación para inferir diseño?. Una cantidad de cosas. Resumiré brevemente lo que está mal punto por punto (para más detalles, refiérase al capítulo dos de *No Free Lunch*).

1. Necesidad de probabilidades a priori. Como ya hemos visto, para que la aproximación bayesiana funcione se requieren probabilidades a priori. Sin embargo, las probabilidades a priori a menudo son imposibles de justificar. A diferencia del ejemplo de la urna y las dos monedas discutidas anteriormente (en el cual extraer una bola de una urna determina netamente las probabilidades a priori de cuál moneda fue lanzada), para la mayoría de las inferencias de diseño, especialmente aquellas que son más interesantes como la existencia de diseño en sistemas biológicos, o no tenemos cómo manipular las probabilidades a priori de una hipótesis de diseño o esa probabilidad a priori es ferozmente discutida (los teístas, por ejemplo, pueden considerar alta esa probabilidad a priori mientras que los ateos la considerarían baja).

2. Hipótesis de diseño que confieren probabilidades. La aproximación bayesiana requiere que las hipótesis de diseño, como las hipótesis de azar, asignen probabilidades a los eventos. En la notación anterior, para que la aproximación bayesiana funcione, las verosimilitudes P(E | D) y P(E | H) deben estar bien definidas. Suponga que E denota el evento responsable de cierto gen, donde este gen a su vez codifica una cierta enzima. Dados los varios procesos naturales a los cuales los genes están sujetos (mutación, supresión, duplicación, cruzamiento, etc), P(E | H) está bien definida. Pero ¿qué pasa con P(E | D)? Asumiendo que la enzima en cuestión constituye una innovación biológica sin precedente ¿cómo asignamos la probabilidad de que un diseñador la diseñe?

Aquí la dificultad no está reducida a las hipótesis de diseño en biología. De hecho, aplica a todos los casos de diseño novedoso. Para estar seguros, hay hipótesis de diseño que confieren probabilidades confiables. Por ejemplo, que yo esté digitando este libro le confiere una probabilidad de cerca del 13% a la letra *e* (ese es el valor de qué tan a menudo un escritor en inglés utiliza la letra *e*). Pero ¿cuál es la probabilidad de que yo escriba este libro?. ¿Cuál es la probabilidad de que Rachmaninoff componga sus variaciones a un tema de Paganini?. ¿Cuál es la probabilidad de que Shakespeare escriba uno de sus sonetos?. Cuando el asunto es sobre creaciones novedosas, el mismo hecho de expresar P(E | D) se torna altamente problemático y prejuiciado. Eso ubica a la creación novedosa de un diseñador en el mismo lugar que las leyes naturales, requiriendo del diseño una predictibilidad que es circunscribible en términos de probabilidades. Pero los diseñadores son inventores de novedades sin precedentes, y tal creación innovativa trasciende todas las probabilidades.

3. La ilusión de rigor matemático. Como hice notar en el punto anterior, si E denota la ocurrencia de cierta codificación genética para una cierta enzima novedosa, entonces se puede considerar que P(E | H) es una probabilidad bien definida. Si el problema de asignar esta probabilidad no es técnicamente muy difícil, puede ser que logremos evaluarla con precisión, o al menos estimar una cota superior para ella. Pero ¿qué pasa con P(E | D)?. ¿Qué pasa con probabilidades como esta, más en general, donde una hipótesis de diseño confiere una probabilidad a una creación novedosa? No sólo no hay razón para pensar que dichas probabilidades tengan sentido (vea el punto anterior), sino que cuando los bayesianos razonan con ellas, lo hacen sin asignarles números precisos. La probabilidad P(E | D) funciona como una representación de la ignorancia, dando un aire de rigor matemático a lo que en realidad es simplemente una asignación subjetiva de qué tan plausible parece una hipótesis de diseño a la persona que ofrece el análisis bayesiano.

4. Eliminación del azar sin comparación. Dentro de la aproximación bayesiana, la evidencia estadística es inherentemente comparativa: no hay evidencia a favor o en contra de una hipótesis como tal sino sólo mejor o peor evidencia a favor de una hipótesis en relación con otra. Pero que todo razonamiento estadístico debe ser comparativo de esta forma no puede ser correcto. Existen casos donde una y sólo una hipótesis estadística es relevante y debe ser determinada. Considere, por ejemplo, una moneda honesta (es decir, un disco rígido perfectamente simétrico con lados distinguibles) que usted mismo está lanzando. Si usted observa mil caras en serie (un evento abrumadoramente improbable), usted estará inclinado a rechazar la única hipótesis de azar relevante, a saber, que los lanzamientos de moneda son independientes e idénticamente distribuidos con probabilidad uniforme.

¿Importa para que rechace esta hipótesis de azar si usted ha formulado una hipótesis alternativa?. Yo digo que no. Para ver esto, pregúntese a usted mismo *¿cuándo empiezo a mirar hipótesis alternativas en tales escenarios?*. La respuesta es, precisamente, cuando un evento altamente improbable como mil caras en serie ocurre. Así, no es que usted haya empezado comparando dos hipótesis sino que empezó con una única hipótesis a la cual, cuando se hizo problemática para dar cuenta de una improbabilidad tan alta (sugiriendo ella misma que las pruebas de significancia fisherianas acechan desde el fondo), usted rechazó tácitamente al inventar una hipótesis alternativa. La hipótesis alternativa en tales escenarios es completamente ex post facto. Es inventada únicamente para mantener viva la ficción bayesiana de que todo el razonamiento estadístico debe ser comparativo.

5. Retroceso de las *a priori*. Como variación del último punto, retornemos al ejemplo anterior de una urna con un millón de bolas, una negra y el resto blancas. Como antes, imagine que una moneda legal se lanza si una bola blanca es seleccionada aleatoriamente de la urna, pero una moneda sesgada con probabilidad 0.9 de caer cara es lanzada en caso contrario. Esta vez, sin embargo, imagine que la moneda se lanza no 10 veces sino diez mil veces y que todas las veces el resultado es cara. La probabilidad de obtener diez mil caras en serie con la moneda legal es aproximadamente 1 en 10^{3010}; con la moneda sesgada e aproximadamente 1 en 10^{458} (con diez mil lanzamientos, es extremadamente probable que resulten sellos con cualquiera de las dos monedas). Un análisis bayesiano muestra entonces que la probabilidad de que se haya seleccionado una bola blanca es aproximadamente 1 en 10^{2546}, y la probabilidad de que la única bola negra sea seleccionada es uno menos esa probabilidad minúscula.

¿Debiéramos por lo tanto, como buenos bayesianos, concluir que la bola negra fue en efecto seleccionada y que la moneda sesgada en efecto fue lanzada? (la selección de la bola negra es grandemente más probable, dadas diez mil caras en serie que la selección de una bola blanca). Esto es claramente absurdo. La probabilidad de obtener diez mil caras en serie con cualquier moneda es grandemente improbable y no importa cuál bola fue seleccionada. La única conclusión sensible es que *ninguna* de las dos monedas fue lanzada aleatoriamente diez mil veces. Un bayesiano puede por lo tanto querer cambiar la probabilidad a priori para introducir duda acerca de si la urna, y subsecuentemente una de las dos monedas, fue aleatoriamente muestreada. Pero, como en el punto anterior, debemos preguntarnos qué nos induce a cambiar o revaluar nuestras probabilidades a priori. La respuesta es que no son estrictamente consideraciones bayesianas sino más bien consideraciones de pequeñas probabilidades basadas en hipótesis de azar que, como ya se dejó claro anteriormente, no admiten alternativa alguna. Las alternativas necesitan, entonces, introducirse subsecuentemente porque consideraciones fisherianas, no bayesianas, las impulsan.

6. Evidencia empírica independiente a favor de diseño. Los teóricos bayesianos están casados a menudo con un marco inductivo como el de Hume, en el cual las hipótesis de diseño requieren evidencia empírica e independiente de un diseñador que efectivamente esté trabajando (es decir, que la cámara esté grabando y el diseñador sea –o al menos en principio pudiera ser- captado en la videocámara) antes de poder atribuir diseño legítimamente. En el capítulo anterior vimos que esta restricción no es sólo artificial sino de hecho incoherente porque la inducción no puede ser la base para identificar diseño, pues no hay forma de obtener y practicar esa inducción. No obstante, para los bayesianos casados con Hume, es conveniente bloquear un análisis bayesiano que pueda implicar incluso empezar a pensar en diseño al negar que ciertas hipótesis de diseño –como una hipótesis de diseño que apele a una inteligencia no evolucionada para explicar la complejidad biológica- podrían aun en principio admitir evidencia empírica independiente.

De este modo, en lugar de enfrentar el problema de asignar probabilidades a priori en tales casos, los bayesianos casados con Hume meramente imponen una restricción adicional en el marco bayesiano al estipular, en efecto, que el marco bayesiano no puede usarse a favor de hipótesis de diseño sin evidencia empírica de un diseñador. Estrictamente hablando, esta restricción no tiene lugar dentro de un aparto probabilístico bayesiano (el teorema de Bayes funciona independiente de donde vengan las probabilidades asociadas con una hipótesis de diseño. Solo inserte los números). Pero dicha restricción se está invocando crecientemente en contra del diseño inteligente. Por ejemplo, mientras Sober permitía considerable libertad a las inferencias bayesianas de diseño en biología en su edición de 1993 de *Philosophy of Biology* [*Filosofía de la*

Biología] (y así antes de que el diseño inteligente tuviera corriente intelectual), en la edición de 2000 del mismo libro cerró la posibilidad de cualquier inferencia de diseño a un diseñador que carezca de evidencia empírica independiente (después de que el diseño inteligente había creado considerables olas). De este modo, mientras la edición de 1993 le dio cabida al diseño inteligente, la edición del 2000 se la quitó.

El requerimiento de evidencia empírica independiente levanta un curioso dilema para el darwinismo. Imagine que se presentan viajeros del espacio cargados con tecnología increíblemente avanzada. Ellos nos dicen (en español) que han tenido esa tecnología por cientos de millones de años y nos dan evidencia sólida de ello (tal vez señalándonos algún aglomerado de estrellas cuyo arreglo significa un mensaje que confirma la afirmación de los extraterrestres). Más aún, nos demuestran que con esa tecnología pueden ensamblar átomo por átomo y molécula por molécula los organismos más complejos. Suponga que tenemos buenas razones para pensar que estos extraterrestres estuvieron aquí en la Tierra en los momentos claves de la historia (por ejemplo, el origen de la vida, el origen de las eucariotas, el origen de los metazoos y el origen del filos animal en el cámbrico). Suponga, aún más, que al formar la vida de la nada los extraterrestres no dejaron ningún rastro (su tecnología es tan avanzada que limpiaron perfectamente todo tras ellos, nada de basura u otras señales de actividad que dejaran tras de sí). Suponga, finalmente, que ninguno de los hechos biológicos es diferente de lo que es ahora. ¿Deberíamos pensar que en momentos claves de la historia la vida fue diseñada?

Ahora tenemos toda la evidencia empírica independiente que podríamos querer para la existencia de diseñadores físicamente corpóreos capaces de producir la complejidad de la vida en la Tierra. Si, adicionalmente, nuestro mejor análisis probabilístico de los sistemas biológicos en cuestión nos dice que los procesos naturales no guiados no pudieron haber producido dichos sistemas con algo semejante a una probabilidad razonable ¿Se garantiza ahora una inferencia de diseño bayesiana? ¿Podría volverse en ese caso el diseño de la vida más probable que una explicación darwinista (siendo aquí interpretadas las probabilidades en un sentido bayesiano o de verosimilitud) simplemente porque la evidencia empírica independiente confirma a diseñadores con la capacidad e producir sistemas biológicos?

Este prospecto, sin embargo, debe preocupar a los darwinistas. Los hechos de la biología, después de todo, no han cambiado. Con todo, el diseño sería una mejor explicación si se pudieran confirmar diseñadores capaces de producir, digamos, el filos animal cámbrico a través de evidencia empírica independiente. Note que aquí no hay ninguna forma de arma caliente (no tenemos evidencia directa alguna de participación extraterrestre en el registro fósil, por ejemplo). Todo lo que por observación sabemos es que existen seres con el poder de generar la vida y que pudieron haber actuado. ¿Nos ayudaría esto a saber que a los extraterrestres en realidad les gusta construir vida basada en el carbono? ¿Pero cómo sabríamos eso? ¿Simplemente creeríamos en su palabra para ello? Los datos de la biología y de la historia natural, suponemos, permanecen como están ahora.

Pero si el diseño es una mejor explicación simplemente por la evidencia empírica independiente de seres extraterrestres tecnológicamente avanzados ¿por qué no debería ser una mejor explicación en la ausencia de tal evidencia? Si el darwinismo es una explicación tan pobre que podría colapsar en el instante que se pudieran verificar independientemente seres extraterrestres capaces de generar formas de vida en toda su complejidad ¿entonces por qué debe dejar de ser una explicación pobre en la ausencia de tales extraterrestres?. Una vez más, los hechos mismos de la biología no han cambiado.

¿Hay alguna forma de salvar el requerimiento de evidencia empírica independiente? Claramente no sería legítimo modificar este requerimiento al descartar la evidencia circunstancial enteramente y permitir sólo evidencia directa de un ”testigo ocular” de un diseñador que en realidad manipule el objeto diseñado en cuestión. Incluso Elliot Sober no seguiría el camino de esta propuesta (vea su *Reconstructing the Past* [*Reconstruyendo el Pasado*]. Para reconstruir el pasado necesitamos evidencia circunstancial). Para Sober, en principio la evidencia circunstancial podría soportar la hipótesis de diseño biológico. Lo importante para Sober es que haya evidencia empírica independiente a favor de la existencia de un diseñador. Pero no se requiere un arma caliente. De hecho, requerir un arma caliente en el sentido de un “testigo ocular” directo sería tan malo para el darwinismo como para el diseño inteligente. La evidencia es simplemente tan circunstancial tanto para el uno como para el otro.

Pero una vez la evidencia empírica independiente a favor de diseño pueda ser circunstancial, estableciendo meramente la existencia de un diseñador con el poder causal y la oportunidad para producir el efecto en cuestión (como en el caso del experimento planteado de los extraterrestres), para poder dar una explicación tenemos exactamente el mismo conjunto de datos biológicos que teníamos antes de que adquiriéramos esa evidencia. Por lo tanto el requerimiento de evidencia empírica independiente es vacío (si esta puede ser circunstancial) o prejuiciado (si se requiere que sea directa). Y en cada caso obstruye la indagación de cualquier forma real de diseño que pueda estar presente. Si requerimos evidencia empírica independiente de diseño pero no la tenemos, no veremos diseño incluso si tal diseño existe.

7. Uso implícito de especificaciones. Y finalmente llegamos al problema más dañino que pueda enfrentar la aproximación bayesiana, a saber, que presupone la misma clase de especificaciones y regiones de rechazo que pretendía excluir. Los teóricos bayesianos ven en las especificaciones una característica dispensable e incongruente de las inferencias de diseño. Por ejemplo, Timothy y Lydia McGrew consideran que las especificaciones no tienen “relevancia epistémica” alguna (Symposium on Design Reasoning [Simposio sobre Razonamiento de Diseño], Calvin College, mayo de 2001). En el mismo simposio Robin Collins, tambien bayesiano, subrayó que “podríamos, a grandes rasgos, definir una especificación como cualquier tipo de patrón para el cual tenemos algunas razones para esperar que un agente inteligente lo produzca”. De este modo el uso bayesiano de la especificación puede verse como sigue: dado algún evento E y una hipótesis de diseño D, una especificación ayudaría a inferir diseño en E si la probabilidad de E condicionada a D crece al notar que E se ajusta a la especificación (la cual, á la Collins, es un “patrón para el cual tenemos algunas razones para esperar que un agente inteligente lo produzca”).

Pero hay aquí una dificultad crucial que los bayesianos invariablemente dejan de lado. Considere el caso del comisionado para las elecciones en New Jersey, Nicholas

Caputo, quien fue acusado de aparejar las líneas en las papeletas de votación (este ejemplo aparece en varios de mis escritos y ha sido ampliamente discutido en Internet. Una línea en la papeleta de votación es el orden en el que los candidatos aparecen listados en la papeleta. Es ventajoso para un candidato estar primero en la lista en una línea de votación porque los votantes tienden a votar más fácilmente para tales candidatos). Llame las selecciones de las líneas de papeleta hechas por Caputo el evento E. E consiste de 41 selecciones de demócratas y republicanos en serie con los demócratas excediendo grandemente en número a los republicanos por 40 a 1. En definitiva, asumamos que las líneas de selección de Caputo en las papeletas se veían como sigue (las noticias que cubrían la historia nunca reportaron la serie real hasta donde yo se):

DDDDDDDDDDDDDDDDDDDDDDRDDDDDDDDDDDDDDDDDD

De este modo, suponemos que las 22 veces iniciales Caputo escogió a los demócratas para encabezar las líneas de las papeletas de votación, luego en la oportunidad 23 escogió a los republicanos, después de lo cual escogió a los demócratas en las oportunidades restantes.

Si los demócratas y los republicanos tuvieron igual probabilidad de encabezar las listas (como afirmó Caputo), este evento tiene probabilidad de 1 en 2 trillones aproximadamente. Esto es improbable, sí, pero en sí mismo no es suficiente para implicar que Caputo está haciendo trampa. Eventos altamente improbables ocurren por azar todo el tiempo después de todo. En realidad, una sucesión cualquiera de 41 demócratas y republicanos sería simplemente igual de inverosímil. ¿Qué necesitamos entonces para confirmar la trampa (y por lo tanto diseño)? Para implicar que Caputo está haciendo trampa, no es suficiente solamente notar la preponderancia de los demócratas sobre los republicanos en alguna sucesión de líneas de selección en las papeletas de votación. En su lugar, se debe notar también que una preponderancia tan extrema como esta es altamente inverosímil. En otras palabras, no era del evento E (las selecciones de la línea en la papeleta de votación como sucedieron) la improbabilidad que el bayesiano necesitaba calcular sino la del evento compuesto E*, el cual consiste de todas las posibles selecciones de línea en las papeletas de votación que exhiben al menos tantos demócratas como seleccionó Caputo. Este evento compuesto E* consiste de 42 posibles selecciones de línea en las papeletas de votación y tiene improbabilidad de 1 en 50 billones. Fue este evento y esta improbabilidad en la cual la Corte Suprema de New Jersey correctamente se enfocó cuando deliberaba si en efecto Caputo había hecho trampa. Más aún, es este el evento que deben identificar los bayesianos y cuya probabilidad deben calcular para hacer el análisis bayesiano.

¿Pero cómo identifican este evento los bayesianos? Seamos claros en que la observación nunca nos da eventos compuestos como E* sino solo resultados elementales como E (es decir, la selección real de la línea en las papeletas de votación de Caputo y no el ensamble de selecciones tan extremas como la de Caputo) ¿Pero de dónde vino este evento compuesto? Dentro del marco fisheriano la respuesta es clara: E* es la región de rechazo (y por lo tanto la especificación) que cuenta el número de demócratas seleccionados en 41 intentos y totaliza al menos tantos demócratas como en la selección de las líneas en las papeletas de votación de Caputo. Eso fue lo que la corte usó y eso es lo que los bayesianos usan. Los bayesianos, sin embargo, no ofrecen ninguna explicación de cómo identificar los eventos a los cuales les asignan probabilidades. Si los únicos eventos que ellos alguna vez consideraran fueran resultados elementales, no habría ningún problema. Pero ese no es el caso. Los bayesianos de manera rutinaria consideran tales eventos compuestos. En el caso de las inferencias bayesianas de diseño (y definitivamente los bayesianos quieren hacer una inferencia de diseño al respecto de las selecciones de las líneas en las papeletas de votación de Caputo), esos eventos compuestos están dados por las especificaciones.

Permítaseme pintar el cuadro más cabalmente. Considere un resultado elemental E. Suponga inicialmente que no vemos patrón alguno que nos de razón para esperar que un agente inteligente produjo el evento. Pero entonces, escudriñando a través de nuestro conocimiento de fondo, repentinamente vemos un patrón que significa el diseño de E. Bajo un análisis bayesiano, la probabilidad de E dada una hipótesis de diseño repentinamente se dispara. Eso, sin embargo, no es suficiente para permitirnos inferir diseño. Como es usual en el esquema bayesiano, necesitamos comparar una probabilidad condicionada a diseño con una condicionada al azar. ¿Pero para cuál evento calculamos estas probabilidades? Como resultan las cosas, no para el resultado elemental E sino para el evento compuesto E* el cual consiste en todos los resultados elementales que exhiben el patrón que significa diseño. En realidad, no hace bien argumentar que E es el resultado de diseño sobre la base de algún patrón a menos que la colección completa de eventos elementales que exhiben ese patrón sea improbable en sí misma bajo la hipótesis de azar. Por lo tanto, el bayesiano debe comparar la probabilidad de E* condicionada a la hipótesis de diseño con la probabilidad de E* condicionada a la hipótesis de azar.

El pié de página es este: la aproximación bayesiana a la racionalidad estadística es parásita de la aproximación fisheriana y puede adjudicar apropiadamente sólo entre hipótesis que la aproximación fisheriana ha fallado en eliminar. En particular, la aproximación bayesiana no da cuenta alguna de cómo llega a los eventos a los cuales les realiza el análisis bayesiano. La selección de esos eventos es altamente intencional, y en el caso de las inferencias bayesianas de diseño debe presuponer una explicación de la especificación. Lejos de ser refutada por la aproximación bayesiana, la complejidad especificada está por lo tanto implícita a través de todas las inferencias bayesianas de diseño.

Para resumir, no hay ningún mérito en la acusación de que al mirar la complejidad especificada para inferir diseño, el diseño inteligente viola la racionalidad estadística. Todo lo contrario. Al desarrollar la complejidad especificada como herramienta analítica para inferir diseño, el diseño inteligente lleva adelante el estudio del razonamiento científico y vindica la aproximación fisheriana de racionalidad estadística.