doi: 10.5007/1808-1711.2014v18n1p115

Raciocínio Científico e Probabilidade: uma comparação entre o bayesianismo e a estatística do erro

Agnaldo Cuoco Portugal

Breno Hermann

Abstract. The article presents two alternative proposals for the use of probability to analyze scientific reasoning: Bayesianism and error statistics. The debate between these two approaches is one of the most important issues in contemporary Philosophy of Science and is a continuation of the well-known debate between Popper and Kuhn. The article presents the explanations offered by Bayesianism for specific phenomena of scientific activity that other approaches have difficulty in explaining, like the ravens paradox. Despite its positive results as a research program, Bayesianism has been the target of strong criticism, for instance, because it allegedly does not offer an adequate solution to Duhemt's problem. Error Statistics in particular proposes the application of statistical methods and probability calculus to explain scientific reasoning in a way radically different from Bayesianism. The debate started by Popper and Kuhn is continued in probabilistic terms and is far from ended.

Keywords: Deborah Mayo; Colin Howson; scientific reasoning; Bayesianism; error statistics.

Introdução

Apesar de decorridos cerca de cinquenta anos, o debate entre Popper (1959) e Kuhn (1962) acerca do papel da experiência na avaliação de teorias científicas e da natureza do raciocínio científico continua fortemente influente na filosofia da ciência atual. As críticas kuhnianas contra a falta de apoio histórico da interpretação popperiana da atividade científica levaram a área a buscar alternativas ao dedutivismo de Popper como forma de descrever o raciocínio em ciências naturais. Embora já desenvolvido desde os anos 1930 e mencionado por Hempel (1966) à época do debate mencionado, a teoria da confirmação de uma hipótese por um indício empírico com base no teorema de Bayes só chamou mais a atenção dos filósofos da ciência no final dos anos 1980, com a primeira edição do trabalho de Howson e Urbach, *Scientific Reasoning – the Bayesian Approach*, publicada em 1989. Começava, então, a promissora trajetória do bayesianismo como metodologia alternativa ao popperianismo, que incorporava elementos subjetivos e não dedutivos na interpretação do raciocínio científico (fazendo jus às alegações de Kuhn), mas que não permitia que estes tornassem relativista e irracionalista a atividade científica teórica (a principal réplica popperiana contra a abordagem do autor de *A Estrutura das Revoluções Científicas*).[1]

Principia **18**(1): 115–134 (2014).
Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

O bayesianismo parecia enterrar de vez a metodologia falsificacionista de Popper e incluir perenemente os elementos probabilísticos na interpretação do raciocínio em ciências naturais. No entanto, como pretende mostrar este artigo, apesar da introdução do cálculo de probabilidade ter sido uma contribuição que perdura ainda hoje, o sentido bayesiano de probabilidade foi alvo de sérias críticas de uma filosofia do raciocínio científico que retomou em vários sentidos o falsificacionismo de Popper, incorporando a ele elementos de probabilidade não bayesiana: a estatística do erro de Deborah Mayo. Vejamos em seguida esses dois sentidos de probabilidade.

Em um famoso artigo na área, Egon Pearson (1950) resume as principais características de duas das interpretações mais conhecidas para a ideia de probabilidade. Segundo a primeira posição descrita por Pearson, chamada de *subjetivista*, a ideia de probabilidade expressa o grau de crença do analista na veracidade de determinada proposição. Trata-se, assim, de um conceito que expressa uma noção essencialmente subjetiva. A segunda posição, chamada de *frequentista*, assume o ponto de vista oposto. Ela entende que a ideia de probabilidade é essencialmente objetiva e pode ser expressa pela frequência, isto é, pela razão do número de ocorrências do evento desejado pelo número total de eventos possíveis.

Na filosofia da ciência, o debate entre subjetivistas e frequentistas deu origem a duas propostas alternativas de aplicação dos métodos probabilísticos para analisar o raciocínio científico. A primeira dessas propostas, o bayesianismo, conta entre seus adeptos nomes da envergadura de Frank Ramsey (1931), Bruno de Finetti (1989) e Harold Jeffreys (1961). O bayesianismo baseia-se na noção subjetivista da probabilidade e faz uso do Teorema de Bayes para melhor compreender o impacto de indícios sobre determinada hipótese.[2] A segunda proposta, que constitui uma resposta ao bayesianismo, é a chamada estatística do erro (*error statistics*), cujo foco de atenção é aplicar procedimentos estatísticos para testar hipóteses de forma rigorosa e, a partir daí, realizar inferências com grau de certeza tão grande quanto possível. Tal proposta, que é herdeira direta do falsificacionismo de Karl Popper, tem na filósofa norte-americana Deborah Mayo sua principal defensora.

Ao longo das últimas décadas, o bayesianismo tem constituído o enfoque dominante na aplicação do cálculo das probabilidades como instrumento para analisar o raciocínio científico. De fato, a abordagem bayesiana foi empregada no exame de algumas das mais importantes questões de filosofia da ciência: ontologia do processo decisório, revisão de crenças, confirmação, indução, natureza da explicação, progresso científico, natureza da crença e do conhecimento científico, entre outras.[3] Na verdade, ele foi apresentado por seus defensores como um sucessor do falsificacionismo popperiano no papel de correta interpretação da lógica da pesquisa científica.

Na segunda metade da década de 1990, porém, críticas ao enfoque bayesiano deram ensejo ao surgimento de enfoque alternativo, formulado por Deborah Mayo e que, ao retomar ideias essenciais ao racionalismo crítico de Popper, agregava a ele o emprego rigoroso de métodos e práticas estatísticas para defender uma proposta que se pretende mais objetiva na análise do raciocínio científico.

Curiosamente, a despeito de sua importância no meio acadêmico norte-americano e europeu, o referido debate permaneceu praticamente ignorado pela comunidade filosófica brasileira.[4] Este artigo se propõe, assim, a ajudar a preencher essa lacuna e apresentar alguns dos principais elementos da discussão entre o bayesianismo e a estatística do erro. Ao fazê-lo, busca identificar os pontos fortes e fracos do bayesianismo e da estatística do erro como propostas rivais de análise do raciocínio científico e continuadoras do debate Popper *versus* Kuhn em filosofia da ciência. Não se pretende fazer uma exposição exaustiva dos dois enfoques, até porque isso seria impossível em um artigo de curta extensão, mas apenas lançar luz sobre alguns dos aspectos principais desse que é, sem dúvida, um dos debates mais relevantes da filosofia da ciência contemporânea.

Este artigo está dividido em seis seções. Na primeira parte, são apresentadas as ideias principais que definem o enfoque bayesiano, conforme descrito no livro *Scientific Reasoning – the Bayesian Approach* (edições de 1989, 1993, 2006), de Colin Howson e Peter Urbach. A seguir, na segunda parte, é apresentada a explicação bayesiana para alguns fenômenos próprios da atividade científica que outros enfoques têm dificuldade em justificar. A terceira parte do artigo se detém sobre a explicação bayesiana para o paradoxo dos corvos. Dada a importância do referido paradoxo, o fato de o bayesianismo ter apresentado uma solução original e satisfatória para o mesmo constitui um elemento de peso em seu favor. Na quarta seção, é apresentada a crítica de Deborah Mayo ao bayesianismo como enfoque incapaz de solucionar adequadamente o problema de Duhem. A partir dessa crítica, distinguem-se as diferenças que separam o falsificacionismo do bayesianismo e torna-se possível divisar as razões que motivaram Mayo a propor o enfoque alternativo da estatística do erro em *Error and the Growth of Experimental Knowledge* (1996). O artigo se encerra com uma comparação entre o bayesianismo e a estatística do erro, ressaltando os pontos de contato e de afastamento entre essas duas propostas de análise do raciocínio científico.

1. O enfoque bayesiano do raciocínio científico

Em 1763, a Sociedade Real de Londres publicou um artigo de um de seus ex-membros, falecido dois anos antes. Tratava-se de "Um Ensaio com o propósito de resolver um problema na Doutrina das Probabilidades" (*An Essay towards Solving a Problem in the Doctrine of Chances*), de autoria de um matemático e pregador não-conformista, o Reverendo Thomas Bayes. Nesse artigo, Bayes não chegou a dar tratamento matemático a suas ideias — tal formalização somente viria a acontecer no século XX —, tendo apenas as apresentado de forma intuitiva. Não obstante, tais ideias vieram a exercer posteriormente uma enorme influência. Uma grande variedade de teorias que carregam o nome de Bayes surgiu em áreas que vão da teoria das probabilidades a modelos de racionalidade científica, teoria da confirmação e inteligência artificial.

A formulação básica do Teorema de Bayes para avaliar a probabilidade de uma determinada hipótese h ser confirmada, tendo em vista o indício e, é a seguinte:

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

Nessa fórmula, afirma-se que a probabilidade posterior da hipótese h, dado um indício e (termo à esquerda do sinal de igualdade), é igual ao quociente da probabilidade subjetiva de e ocorrer, dada a hipótese h, pela probabilidade prévia de e $\frac{P(e/h)}{P(e)}$ multiplicado pela probabilidade prévia — isto é, independente da ocorrência de e — da hipótese h $[P(h)]$.

A premissa de base para a aplicação do Teorema de Bayes ao raciocínio científico é a constatação de que as informações obtidas por meio da observação exercem um papel essencial para confirmar ou infirmar uma hipótese. No enfoque bayesiano, o aprendizado a partir da experiência é, essencialmente, uma questão de grau mensurável probabilisticamente. Assim, poderíamos dizer que (Howson & Urbach 2006, p.92):

(a) e confirma ou apoia h caso $P(h/e) > P(h)$, isto é, a probabilidade posterior da hipótese é maior do que sua probabilidade prévia, uma vez tendo se verificado o indício e;

(b) e infirma h caso $P(h/e) < P(h)$, isto é, a probabilidade posterior da hipótese é menor do que a probabilidade prévia, uma vez tendo se verificado o indício e;

(c) e é neutro com relação a h caso $P(h/e) = P(h)$, isto é, a ocorrência do indício e não afeta a probabilidade da hipótese h;

Tais asserções, que inicialmente podem parecer simples definições, constituem, na verdade, o núcleo do argumento bayesiano. Uma vez atribuídos valores a $P(e/h)$, $P(e)$ e $P(h)$, é possível determinar se e confirma ou não h, ou seja, qual é o valor de $P(h/e)$. Howson indica que, na prática, é comum que as probabilidades sejam conhecidas apenas de forma imprecisa, mas isso não inviabiliza a aplicação do Teorema de Bayes como forma de chegar a uma justificativa para a inferência científica. Para dadas probabilidades prévias $P(e)$ e $P(h)$, diz-se que e refuta h caso a probabilidade posterior de h, dado e, seja igual a zero, isto é, $P(e/h) = 0$. Nessa situação, o grau de infirmação é máximo. De forma semelhante, o grau máximo de confirmação ocorre quando a probabilidade posterior de h, dado e, for igual a 1, isto é, $P(e/h) = 1$, o que se dá quando e decorrer logicamente de h. Ao estabelecer valores para $P(e/h)$, estamos adstritos aos axiomas do cálculo de probabilidades, um dos quais estabelece que a probabilidade de um evento deve ser expressa por um número entre 0 e 1.

Na teoria subjetivista das probabilidades, que é a adotada por Howson e Urbach,[5] o um valor da probabilidade prévia de uma hipótese h $[P(h)]$ é o grau de crença que um cientista tem na verdade desta, considerando o que ele sabe pessoalmente sobre ela e o assunto a que esta se refere. Dessa forma, a atribuição de determinado valor a $P(h)$ expressa menos a probabilidade objetiva de veracidade da hipótese, do que o quão provável o cientista acredita que a hipótese seja. Uma interpretação possível para isso é a de Ramsey, para quem a probabilidade prévia deve ser interpretada como a "atitude" do cientista com relação à hipótese, a qual pode ser medida por meio do montante que este estaria disposto a apostar em sua veracidade.

Antes de analisarmos críticas ao raciocínio científico baseado nessa concepção de probabilidade, vejamos alguns resultados bem sucedidos do bayesianismo.

2. Algumas virtudes bayesianas

Uma boa parte da atração do enfoque bayesiano pode ser atribuída a seu êxito em explicar princípios e práticas da metodologia científica a partir de um teorema não controverso do cálculo de probabilidades, de premissas aceitáveis acerca do comportamento racional do agente e da postulação de mudanças ordenadas no grau de expectativas do investigador.[6]

É possível, por exemplo, entender por que razões determinado indício confirma uma hipótese h com especial força se tal indício for particularmente inesperado ou surpreendente. Tomem-se as seguintes formulações equivalentes do Teorema de Bayes (Howson e Urbach 2006, p.97):

$$P(h/e) = \frac{P(e/h)}{P(e)}) \times P(h) = \frac{1}{P(h) + P(\sim h) \times \frac{P(e/\sim h)}{P(e/h)}}.$$

No caso de e ser uma consequência de h, em que $P(e/h) = 1$, o grau de confirmação de h é inversamente proporcional a $P(e)$, conforme a formulação à esquerda do sinal de igualdade, ou a $P(e/\sim h)$, na formulação à direita. Assim, quanto menor for a probabilidade prévia $P(e)$, ou seja, quão mais improvável for e sem considerar a hipótese h, maior será a confirmação com relação a h. Da mesma forma, quanto menor for a probabilidade de e dado que h seja uma hipótese falsa, isto é, $P(e/\sim h)$, maior será a probabilidade posterior de h. Na verdade, $P(e)$ relaciona-se com as probabilidades de a hipótese h ser verdadeira ou falsa pela seguinte fórmula: $P(e) = P(e/h) \times P(h) + P(e/\sim h) \times P(\sim h)$ (Howson & Urbach 2006, p.19). Notese, por fim, que a probabilidade posterior da hipótese h depende da probabilidade prévia da própria hipótese, isto é, $P(h)$.

O bayesianismo é plenamente compatível com a possibilidade de uma teoria ser refutada pela experiência, conforme exige a concepção popperiana de ciência. Se uma hipótese h implica e como consequência, então $P(e/h) = 1$ e $P(h/{\sim} e) = 0$, desde que $P(h) > 0$. No jargão bayesiano, isso significa que h é infirmada em grau máximo quando é refutada. Tal refutação persistirá caso novos indícios similares a e sejam coletadas. Se uma nova evidencia e' for logicamente consistente com e, teremos que $P(h/{\sim} e\&e') = 0$. De forma análoga, o teorema também mostra que uma teoria é confirmada por todo indício que seja dela decorrente.

Por outro lado, com base no teorema de Bayes, é possível oferecer explicações plausíveis para práticas da atividade científica em que o modelo falsificacionista de Popper se encaixa apenas com dificuldade. Por exemplo, o emprego por este último da lógica dedutiva como forma de interpretar o raciocínio científico acaba não sendo diretamente aplicável a teorias científicas probabilísticas.[7] Além disso, na metodologia popperiana, é prescrita a atividade de teste o mais severo e contínuo possível das conjecturas científicas. O problema é que não há no racionalismo crítico de Popper uma razão para justificar por que, a partir de determinado ponto, não tem mais cabimento continuar a realizar os mesmos experimentos. Porém, é possível mostrar que, ao repetir infinitamente o mesmo experimento, a partir de certo ponto, a sequência de probabilidades posteriores chega a um limite. Isso significa que

$$\lim_{n\to\infty} P\left(\frac{e}{h_1 \ldots \&e_n}\right) = \lim_{n\to\infty} P\left(\frac{e}{h_1 \ldots \&e_{n-1}}\right).$$

Ao estender o número de experimentos ao infinito, o limite das probabilidades posteriores tende a 1, o que mostra a falta de sentido de continuar a repetir a mesma experiência indefinidamente (Howson & Urbach 2006, p.94). É verdade, porém, que o teorema não permite saber a partir de que ponto as repetições se tornam desnecessárias, pois isso exigiria conhecer a sequência de probabilidades posteriores que o cientista atribui a cada repetição, o que é impossível. Em termos bayesianos, contudo, o problema fica claro da seguinte maneira: num experimento igual, a probabilidade do indício e em vista de h $[P(e/h)]$ não se altera — tornando inútil a continuação do mesmo experimento em relação à hipótese em questão —, fazendo diferença apenas testar o quão e é improvável em vista da negação de h $[P(e/\sim h)]$, ou seja, cabendo apenas testar outras hipóteses rivais para explicar e e comparar os respectivos graus de confirmação.

A confirmação bayesiana explica, ainda, as razões pelas quais uma versão restrita h' de uma teoria h é capaz de confirmar h apenas até certo ponto. Tome-se, por exemplo, para confirmar a Teoria de Newton, uma versão simplificada que explique a queda livre dos corpos. Uma vez que a versão restrita h' é uma decorrência lógica da versão completa, então $P(h) \leq P(h')$ (Howson & Urbach 2006, p.17). A partir da acumulação sucessiva de indícios decorrentes de h' (que também são decorrências de h), é possível deduzir que o limite da probabilidade posterior de h é dado pelo fator $P(h)/P(h')$.[8] Isso significa que a probabilidade posterior de h nunca poderá ser maior do que a razão $P(h)/P(h')$. Com isso, a probabilidade prévia de h' $[P(h')]$, que aparece no denominador, funciona como um limite para a probabilidade posterior de h. O que isso significa do ponto de vista da prática científica? Significa que a versão mais reduzida h' confirma a hipótese original h apenas até certo ponto e o faz tanto menos quanto maior for o valor da probabilidade prévia $P(h')$. O fato de uma versão mais reduzida de uma teoria ou hipótese apenas confirmar esta até certo ponto constitui uma constatação comum no dia a dia dos cientistas. Trata-se, assim, de mais uma situação que o bayesianismo consegue explicar a contento.

3. O paradoxo dos corvos

A esta altura, está claro que, para o bayesianismo, a confirmação é uma questão de grau. Isso fica muito evidente a partir da explicação que o bayesianismo ensaia para o chamado paradoxo dos corvos ou paradoxo da confirmação, proposto por Hempel, em 1945, no artigo *Studies in the Logic of Confirmation*. Hempel descreve o paradoxo por meio da hipótese:

(1) Todos os corvos são pretos;
Em termos lógicos, essa hipótese é equivalente à sua contraposição, isto é:

(2) Tudo o que não é preto não é um corvo.

Em todas as circunstâncias em que (2) é verdadeira, (1) também o é. De forma similar, em todas as circunstâncias em que a afirmativa (2) for falsa, a afirmativa (1) também será falsa. É evidente que a observação de um corvo preto confirma a hipótese (1), assim como sua equivalente (2). O paradoxo se verifica ao se estar diante de uma instância que confirma (2), como, por exemplo, quando se observa uma maçã verde. Pelo raciocínio anterior, a observação da maçã verde, ao confirmar (2), também confirma (1), o que é contraintuitivo.

Duas condições específicas são exigidas pelo paradoxo:

(1) Hipóteses do gênero "todos os corvos são pretos" são confirmadas pela observação de um objeto que seja, ao mesmo tempo, "corvo" e "preto" (*condição de Nicod*);

(2) Hipóteses equivalentes do ponto de vista lógico são confirmadas pelo mesmo indício (*condição de equivalência ou condição de Hempel*);

É possível tentar solucionar o paradoxo com base na premissa de que as duas condições não são mutuamente consistentes. Assim, uma das formas de resolvê-lo seria pelo abandono de uma das duas condições identificadas acima ou por sua relativização. Esse não foi, porém, o caminho seguido por Hempel, que preferiu manter a validade das duas condições e, com isso, aceitar o resultado de que a observação de uma maçã verde confirma a hipótese de que todos os corvos são negros. Na verdade, Hempel considerou que o paradoxo era apenas aparente, em função de possuirmos informações prévias sem as quais não haveria qualquer desconforto no fato de uma maçã verde constituir um indício confirmador de que todos os corvos são pretos.

Por que Howson e Urbach analisam o paradoxo dos corvos? Isso acontece porque eles querem não apenas mostrar que o bayesianismo explica inúmeras questões complexas da filosofia da ciência, mas também que o faz melhor do que propostas teóricas alternativas. Particularmente, o bayesianismo pretende se sair melhor na resolução do paradoxo dos corvos do que a interpretação do raciocínio científico baseada na lógica dedutiva apenas, como é o caso do falsificacionismo popperiano.

Ao explicar o paradoxo dos corvos com base no Teorema de Bayes, Howson e Urbach argumentam que se trata de uma situação menos problemática do que inicialmente se supõe. Na solução bayesiana, também são conservadas a condição de equivalência e a condição de Nicod, muito embora esta última não seja, para Howson e Urbach, um princípio de confirmação universalmente válido.

Seja CP a expressão de um objeto "corvo" e "preto" e $\sim C \sim P$ a expressão de um objeto que não é "corvo", nem "preto". Seja, ainda, Θ a expressão dos valores possíveis para expressar a proporção dos corvos que são pretos (nossa hipótese h diz, precisamente, que $\Theta = 1$). É possível mostrar (Howson & Urbach 2006, p.101) que:

$$(1) \quad \frac{P(h/CP)}{P(h)} = \frac{1}{\sum \Theta P(\Theta)} \qquad\qquad (2) \quad \frac{P(h/{\sim}C{\sim}P)}{P(h)} = \frac{1}{P({\sim}C/{\sim}P)}$$

Segundo a equação (1), a razão da probabilidade posterior pela probabilidade prévia de h é inversamente proporcional ao fator $\sum \Theta P(\Theta)$. Isso significa que, se a probabilidade prévia de todos os corvos serem pretos for alta, a observação de um corvo preto terá impacto para a confirmação da hipótese h, mas este será pequeno. Se, por outro lado, a probabilidade prévia de que corvos não são pretos for alta, o grau de confirmação daquela observação será considerável.

A equação (2), por sua vez, diz respeito à confirmação exercida pela observação de um objeto que não é corvo nem preto. Nesse caso, a razão da probabilidade posterior pela probabilidade prévia de h é inversamente proporcional ao fator $P(\sim C/\sim P)$. Nesse caso, como a probabilidade de observação de um objeto com essas características é alta, o impacto sobre a probabilidade posterior será muito pequeno, o que confirma nossa intuição sobre o assunto. Assim, é possível compreender, por meio da explicação bayesiana, porque, intuitivamente, a observação de um corvo preto confirma muito mais a hipótese “todos os corvos são pretos” do que a observação de uma maçã verde. Na solução bayesiana para o paradoxo dos corvos, foram conservadas como válidas tanto a condição de equivalência, quanto a condição de Nicod. A condição de equivalência deve, porém, ser ponderada: embora um indício e confirme tanto a hipótese original quanto outra logicamente equivalente a ela, o efeito de confirmação será diferenciado em um caso e noutro, o que é intuitivamente correto.

4. Deborah Mayo, o problema de Duhem e o bayesianismo

Uma das principais críticas do bayesianismo a Popper é a incapacidade deste de distinguir se um enunciado básico refuta a hipótese principal ou alguma hipótese auxiliar. Howson e Urbach argumentaram que a abordagem por eles defendida era capaz de lidar com este que é comumente chamado “o problema de Duhem” (Howson & Urbach 2006, p.113).

Contudo, em um artigo de 1997, intitulado *Response to Howson and Laudan*, Deborah Mayo defende que o bayesianismo é que é incapaz de solucionar o problema de Duhem. Ao apontar tais falhas, Mayo não está interessada apenas em exibir as insuficiências do bayesianismo, mas em trazer à luz determinadas características desse enfoque que, no seu entender, justificam a postulação de proposta teórica alternativa para aplicar o cálculo de probabilidades na análise do raciocínio científico.

Um primeiro problema que se coloca é que, no enfoque bayesiano, para solucionar o problema de Duhem, o cientista necessita saber a probabilidade condicional de a hipótese h ser falsa, dado que é aceita. Suponha-se um “newtoniano incorrigível” que aceita a hipótese A': o efeito de deflexão da luz é devido a um fator N qualquer (por exemplo, um efeito próprio das lentes de observação) que permite salvar as leis de Newton da refutação. Sabe-se, hoje, em função da Teoria da Relatividade, que A' é falsa. Nessas condições, o que interessa a um cientista confrontado com determinada instância de refutação é a probabilidade condicional de A' ser falsa, dado que é aceita. Ocorre que tal probabilidade condicional não é fornecida pelo bayesianismo, que apenas pode fornecer a probabilidade subjetiva para o agente de que A' é falsa, dado que é aceita. Em outras palavras, o bayesianismo não permite avaliar a veracidade ou falsidade da hipótese, dado que ela é aceita. Ele apenas fornece o grau subjetivo posterior de crença de que a hipótese é verdadeira ou falsa, dado que é aceita (Mayo 1997a, p.325–326).

O cerne da questão enfatizada por Mayo é que o bayesianismo se equivoca ao atribuir um papel central ao grau de crença do cientista em determinada hipótese. Na verdade, o bayesianismo não está preocupado primordialmente com a veracidade da crença, mas sim se tal crença é alcançada segundo um processo consistente e coerente. O que deveria importar, segundo Mayo, é saber as razões que permitem justificar a crença na veracidade ou falsidade de determinada hipótese de maneira objetiva. Ela acredita, portanto, que o bayesianismo não dá ao cientista aquilo que ele mais deseja, isto é, condições de posicionar-se quanto à veracidade ou falsidade de uma hipótese com base em critérios objetivos.

Com vistas a tornar mais clara essa diferença de enfoque entre o bayesianismo e a estatística do erro, Mayo fornece um exemplo em que contrasta os dois enfoques e aponta para as deficiências, segundo ela insuperáveis, do bayesianismo. O exemplo pode ser apresentado de maneira resumida da seguinte forma: imagine-se certa população que apresenta determinada deficiência com a probabilidade de 0.999. Qualquer indivíduo escolhido aleatoriamente dessa população apresenta, portanto, uma probabilidade prévia igual a 0.999 de exibir a referida deficiência (por exemplo, a incapacidade para ser aprovado(a) em um exame competitivo para as melhores universidades do país). Considerem-se, diante dessa situação hipotética, duas hipóteses alternativas:

H: o (a) estudante é aprovado no exame — ele(a) não apresenta a deficiência;

H': o (a) estudante não é aprovado no exame — ele(a) apresenta a deficiência.

Há dois resultados passíveis de observação nesse caso: e (aprovação no exame) e e' (não aprovação no exame). Suponha-se, ainda, que os estudantes que não apresentam a deficiência indicada são aprovados no exame com probabilidade prévia igual a 1; os estudantes deficientes são aprovados no exame com uma probabilidade prévia igual a $0,05$.

Com base nas probabilidades atribuídas, Mayo agora quer mostrar que o bayesianismo leva a um resultado que tem grandes chances de ser falso. A candidata Mary foi aprovada no exame competitivo para a universidade. Diante desse fato, qual das hipóteses deveria um cientista racionalmente escolher, H ou H'? Segundo a leitura que Mayo faz de Howson, um cientista bayesiano estaria obrigado a escolher a segunda hipótese, pois a probabilidade posterior de H' será necessariamente maior do que a probabilidade posterior de H. O fato de a estudante ter sido aprovada terá um impacto muito maior em H' do que em H. De fato, com os dados estipulados, tem-se que $P(H'/e) = 0,98$. O resultado observado confirma muito mais a hipótese de que a estudante apresenta a deficiência do que a hipótese contrária. Isso acontece porque a estudante foi escolhida em uma área em que um indivíduo selecionado aleatoriamente tem probabilidade prévia igual a $0,999$ de não estar em condições de ser aprovado no exame. A aprovação de Mary constitui um evento inesperado com relação a H' e, assim, exerce um efeito de confirmação considerável sobre aquela hipótese, o que explica sua probabilidade posterior ser alta.

Mayo esclarece que o problema com o enfoque bayesiano no exemplo acima é que em momento algum foi considerada a possibilidade de H ser rejeitada erroneamente. E isso é o que interessaria aos cientistas, pois o fato de Mary ter sido aprovada no exame não garante que ela pertencia ao grupo dos deficientes, por mais que essa hipótese tenha sido confirmada em grau muito maior do que a sua rival na ótica bayesiana. Muito provavelmente, ela pertence ao grupo dos que não apresentam a deficiência e a hipótese H foi rejeitada erroneamente pela aplicação do enfoque bayesiano.

O cientista devoto do emprego dos métodos estatísticos não rejeitará H ao observar e, nem tampouco acreditará que esse indício é razão para aceitar H'. Caso a estudante Mary tivesse sido selecionada ao acaso de outra amostra populacional, em que as probabilidades prévias relativas à deficiência fossem distintas — por exemplo, em que a probabilidade prévia de ter a deficiência fosse $0,5$ e não $0,9$ —, então o impacto da aprovação sobre a hipótese H' seria muito menor. Com isso, o mesmo resultado, que em determinada situação levaria o analista a concluir, segundo o enfoque bayesiano, pela deficiência de Mary, no segundo exemplo levaria à conclusão oposta. Para Howson e Urbach, isso seria uma consequência natural da modificação das probabilidades prévias e um resultado coerente do ponto de vista bayesiano.

O problema com o bayesianismo, segundo Deborah Mayo (1997a, p.328–329), é que embora ele seja internamente consistente, ele não dá ao cientista aquilo que ele de fato necessita. No caso de Mary, o que importa é garantir que o indício seja interpretado de forma correta, o que significa excluir qualquer erro de interpretação relativo ao desempenho da candidata. O objetivo a ser buscado, assim, é avaliar a possibilidade de ser erroneamente selecionada a hipótese de Mary ser deficiente. E isso o bayesianismo, não consegue.

5. A estatística do erro

Em contraste com o bayesianismo, os métodos e modelos clássicos da estatística, como os testes de significância e os intervalos de confiança, constituem exemplos da chamada estatística do erro. Tais métodos e modelos rejeitam a possibilidade de atribuição de probabilidades prévias, a menos que estejam baseadas em frequências objetivas. A atribuição de probabilidades, por sua vez, é um instrumento para escrutinar e investigar o processo experimental. Todo processo experimental é caracterizado por margens de erro que podem ser expressas por probabilidades, as chamadas probabilidades de erro.

A estatística do erro apresenta duas dimensões complementares: uma filosófica e outra metodológica (Mayo 2010). A dimensão filosófica diz respeito a um enfoque próprio para a filosofia da ciência, que se baseia no papel do cálculo das probabilidades para realizar inferências. A dimensão metodológica, por sua vez, diz respeito ao emprego de um conjunto de métodos e procedimentos estatísticos, sua interpretação e justificação. Em última instância, a estatística do erro busca oferecer respostas a duas questões, que constituem seu foco principal: a) Como obter conhecimento confiável sobre o mundo, apesar da existência de incertezas e erros?; e b) Qual o papel das probabilidades na realização de inferências?

Na prática do dia a dia, há várias situações em que os métodos contemplados pela estatística do erro são empregados. É o caso, por exemplo, de pesquisas de opinião em que se procura inferir, com margem de p% a mais ou a menos, a proporção de um grupo que, provavelmente, irá votar no candidato X. Outros exemplos são relatórios sobre valores estatisticamente relevantes das diferenças entre um grupo tratado com um medicamento e seu respectivo grupo de controle; e análises de informação em experimentos em Física com o propósito de distinguir efeitos reais de artefatos. Em todos esses exemplos, a noção de erro experimental é essencial. Apesar desse uso difundido, a influência bayesiana fez, segundo Mayo, que por muito tempo os métodos aludidos tenham sido considerados inválidos pelos filósofos da ciência.

É preciso fazer um esclarecimento acerca do exato âmbito da estatística do erro. Embora esse enfoque faça uso de métodos típicos da estatística clássica, inclusive testes de significância e de Neyman–Pearson, ela o faz com vistas a constituir um enfoque novo na interpretação da lógica da pesquisa científica. Tais métodos não constituem um fim em si mesmo. A cada experimento, os testes são empregados para permitir que uma hipótese seja testada da forma mais severa possível. A aplicação de testes de significância, segundo a estatística do erro, não se presta à aceitação ou descarte automático de hipóteses. Cada aplicação do teste deve ser revista à luz das especificidades do experimento em questão, tendo em vista o objetivo maior de realizar inferências confiáveis. O cientista que emprega a estatística do erro em sua prática laboratorial será levado a realizar questionamentos que vão muito além da mera aplicação do receituário recomendado nos manuais de estatística (Mayo 1997b, p.196–198).

A estatística do erro procura desenvolver instrumentos para lidar com as inevitáveis limitações e erros próprios da atividade experimental. Diferentemente do bayesianismo, ela não propõe nem dispõe de um mecanismo automático para avaliar o impacto de indícios em hipóteses. Os métodos estatísticos são empregados pelos investigadores para evitar que suas conclusões sejam equivocadas. Eles permitem aos cientistas desenvolver estratégias para coletar e modelar indícios, verificar a ocorrência de uma variedade de erros e, enfim, realizar inferências com o máximo grau de certeza possível.

Há dois aspectos da estatística do erro que são especificamente ressaltados por Mayo como virtudes em comparação com o bayesianismo (Mayo 1997b, p.199). Em primeiro lugar, em lugar de fazer uso direto de indícios, a estatística do erro chama a atenção para a maneira segundo a qual estes foram recolhidos e modelados. Com frequência, será necessário realizar inferências específicas para que se possa chegar a indícios corretos do ponto de vista experimental (por exemplo, a necessidade de padronizar uma série de medidas para poder interpretá-las corretamente). Em segundo lugar, a estatística do erro não considera que a realização de inferências científicas possa ser levada a cabo através da aplicação de uma fórmula, como o Teorema de Bayes.

A estatística do erro constitui uma moldura investigativa (*a framework of inquiry*) para avalizar a realização de inferências. Não é possível simplesmente confrontar o cientista adepto da estatística do erro com determinado conjunto de indícios e esperar que ele esteja em condições de indicar, de imediato, se e qual hipótese é confirmada por esse conjunto. A moldura investigativa que propõe a estatística do erro exige que o procedimento para responder a tais perguntas seja fragmentado em uma série de etapas (*piecemeal approach*). Para se chegar à conclusão desejada, por exemplo, se os indícios confirmam ou não a hipótese, há uma sequência de passos a ser observada. Esses passos dizem respeito à realização de testes e identificação de possíveis erros desde a concepção de experimentos até sua efetiva realização, passando pela geração e modelagem de indícios.

Assim, de acordo com esse enfoque, uma investigação científica deve ser dividida em hipóteses separadas, as quais devem ser objeto de investigações independentes. Tais investigações independentes podem dizer respeito, por exemplo, a estimativas acerca de valores previstos pela teoria ou a testes hipotéticos dos mesmos. Elas correspondem a questões expressas em termos de erros-padrão de parâmetros, causas, efeitos acidentais e premissas envolvidas na investigação de outros erros. Deve-se identificar determinado erro e, a partir dele, inferir qual hipótese ou qual aspecto de uma hipótese é experimentalmente avalizado. Os modelos experimentais constituem o elemento de ligação entre modelos primários (a teoria) e os indícios modelados conforme o experimento.

Há, portanto, diferentes situações em que a utilização dos métodos estatísticos pode se dar. Uma primeira situação é quando esses métodos são empregados para analisar inferências a partir de hipóteses experimentais. Uma segunda situação é empregá-los para relacionar as inferências obtidas das hipóteses experimentais às previsões da teoria. Uma terceira situação diz respeito, por fim, à análise da transformação dos indícios de seu estado original em indícios modelados, isto é, que estão em condições de serem utilizados nas inferências a partir das hipóteses experimentais. A inferência de uma hipótese a partir de indícios pode ser descartada em função de um erro identificado na hipótese experimental ou porque os indícios obtidos não satisfazem ao que era previsto por essa mesma hipótese experimental.

As considerações acima permitem entender porque, para Mayo (1997b, p.200), não é possível aspirar realisticamente às virtudes que o bayesianismo prega. No mundo real, os cientistas são confrontados com situações em que os indícios são inexatos, incompletos ou em que fatores externos não foram controlados ou não são controláveis. Inferências realizadas sem qualquer tipo de controle levarão inevitavelmente a resultados equivocados. Além disso, os cientistas não têm à sua disposição uma lista exaustiva de hipóteses e suas respectivas probabilidades prévias. Diferentemente do que propõe o bayesianismo, uma hipótese não deve ser aceita porque os indivíduos creem nela firmemente, nem é possível esperar até que eventuais desavenças acerca do poder de confirmação ou infirmação de determinado indício sejam decididas.

Ao utilizar a estatística do erro, o cientista pode descobrir muito acerca do mundo sem precisar da enorme quantidade de informação que, alegadamente, o bayesianismo exige, como a atribuição exaustiva de probabilidades prévias a hipóteses rivais. Mesmo em situações em que a condução de experimentos exaustivos não seja possível, Mayo indica que sempre será possível recorrer a experimentos contrafactuais ou realizar experimentos parciais de forma sequencial.

Uma questão que se coloca, a esta altura, é saber como é possível aprender com os erros. De forma simplificada, a busca pela identificação de erros nas várias etapas de uma investigação científica leva à elaboração de um conjunto de critérios eficientes para detectá-los. Se a despeito de todos os esforços empreendidos nenhum erro for encontrado, então o cientista tem excelentes motivos para acreditar que, de fato, nenhum erro está presente (Mayo 1996, p.202–204). Ou seja, em termos popperianos, a hipótese está corroborada após a realização de um teste rigoroso. Tal conclusão é resultado do emprego de procedimentos para testar de forma estrita a existência de determinado erro e mensurar seu valor (*severely probing error*).

Uma estratégia complementar identificada por Mayo (2011) para se aprender com os erros é argumentar a partir deles (*arguing from error*). Após identificar e conhecer o suficiente sobre determinados erros próprios da prática experimental, o cientista está em condições de montar um experimento capaz de demonstrar a presença de um erro específico, caso este exista.

Um erro é considerado ausente na medida em que determinado procedimento investigativo, que pode ser composto por um ou vários experimentos, tem uma probabilidade alta de identificar esse erro se ele de fato estiver presente e, nessas condições, não o detecta. Do ponto de vista de uma hipótese H que se quer testar, o raciocínio proposto pela estatística do erro funciona da seguinte forma: indícios conformes à hipótese H indicam que ela é correta apenas e na exata medida em que tais indícios são o resultado de um procedimento experimental com alta probabilidade de produzir resultados não tão de acordo com H, caso esta seja falsa. Assim, os indícios indicam que H é correta se essa hipótese for submetida a um teste severo, no qual ela muito provavelmente não passaria se fosse falsa. Com essa concepção metodológica, Mayo preserva a essência do falsificacionismo de Popper, mas inclui algo que ele não admitia: uma heurística, ou seja, um conjunto de procedimentos destinados a selecionar que hipótese adotar para, só então, submetê-la a testes.

A fim de ilustrar o enfoque preconizado pela Estatística do Erro, tome-se uma hipótese H que estabelece que determinado efeito é real. A hipótese H estabelece, por exemplo, que existe uma correlação entre um gene específico e determinado tipo de câncer em uma população. Suponha, agora, que testes experimentais indicam a veracidade da hipótese H. Uma das conclusões passíveis de serem extraídas da veracidade de H é que a incidência do referido gene entre os pacientes que tem câncer não deverá variar da mesma maneira que se esperaria se essa correlação fosse aleatória. Nesse contexto, pode-se considerar a postulação de H como bem sucedida se ela for significativa do ponto de vista estatístico de forma distinta do que se poderia esperar se a relação entre o gene e a doença fosse aleatória. No caso de os testes indicarem que H é falsa, então resultados significativos estatisticamente acima de um nível determinado α, correspondente à correlação aleatória, não ocorrerão ou serão raros.

A noção de severidade se refere sempre a uma inferência em particular ou a uma hipótese que foi testada. Um experimento pode ser severo para testar uma hipótese, mas não outra. Nos parágrafos precedentes, a noção de severidade foi identificada com alta probabilidade de detectar o erro, se esse existir. Não há, porém, necessidade de que a noção de severidade seja construída com relação a qualquer cálculo estatístico. Com frequência, argumentos a partir do erro são construídos em avaliações qualitativas da noção de severidade (Mayo 2011, p.162).

Para testar de forma estrita a existência de determinado erro, o cientista deve subdividir a investigação que pretende levar a cabo em uma série de questões menores e relacionadas entre si. A cada questão corresponde uma hipótese parcial sobre a existência de um erro particular. A afirmação de que a hipótese H é verdadeira ou falsa dependerá da identificação de erros específicos. Por exemplo, se H afirma que determinado erro está ausente, H será falsa se esse erro estiver presente. Se H é verdadeira se um parâmetro qualquer for maior que c, então H será falsa se esse parâmetro for menor que c. Se H afirma que um fator x é responsável por ao menos p% de um efeito, sua negativa afirmará que esse fator é responsável por menos do que p%. Com isso permite-se que cada uma das questões particulares em que determinada investigação foi subdividida seja examinada de forma isolada.

Ao avaliar se uma pessoa tem dor de garganta causada pelo vírus x, o cientista adepto da estatística do erro deverá realizar um ou vários testes experimentais para discriminar o vírus x de outras possíveis causas. Ele trabalha com a dicotomia vírus x *versus* outras causas. Ele não está preocupado em assignar probabilidades prévias às várias hipóteses que poderiam explicar a dor de garganta. No bayesianismo, por outro lado, para avaliar uma hipótese singular é necessário atribuir probabilidades prévias a todas as hipóteses concorrentes. Isso ocorre, segundo Mayo, porque o bayesianismo trabalha com noção de um valor fixo de probabilidade total, igual a um, que deve ser distribuído entre as várias hipóteses rivais (Mayo 1997b, p.204).

Chegamos aqui a um ponto ideal para finalizar o presente artigo. Trata-se de um excelente exemplo de que este é um debate ainda em curso e que toca em alguns dos pontos mais importantes da Filosofia da Ciência como área do conhecimento. Em resposta à crítica acima, os bayesianos podem simplesmente afirmar que a atribuição de probabilidades prévias às hipóteses relevantes não é assim algo tão trabalhoso, uma vez que eles trabalham com a tese de que os cientistas partem de suas crenças subjetivas na consideração das diferentes hipóteses. Essas crenças, como já havia ensinado Kuhn, não são idiossincráticas e meramente individuais, mas o resultado do treinamento recebido na formação e na participação em uma comunidade investigativa.[9] Assim, não custa nada à metodologia bayesiana admitir que um cientista vai atribuir diferentes probabilidades prévias às hipóteses em questão, dependendo do quanto estas são plausíveis dado o conhecimento partilhado pela comunidade científica e do quanto ele pessoalmente acredita que uma hipótese tem mais potencial explicativo do que outras. Isso significa também que hipóteses não consideradas pela comunidade e que não tenham plausibilidade alguma para o pesquisador podem ser simplesmente ignoradas, ou seja, ter probabilidade prévia igual a zero.

É claro que aqui cabe novamente a constante crítica de falta de objetividade feita à abordagem bayesiana e que é também a principal objeção à imagem kuhniana das ciências empíricas. Como vimos, o bayesianismo sustenta que algum elemento de subjetividade é inevitável no raciocínio científico (e nisso Kuhn estava certo) e que sua proposta tem a vantagem de deixar claros os pressupostos assumidos. Por outro lado, o Teorema de Bayes mostra como esse elemento subjetivo pode ser superado pela confirmação ou infirmação com base nos indícios empíricos (o que responde ao problema da imagem kuhniana de ciência acima mencionada). Em outras palavras, segundo Howson e Urbach,[10]

> . . . o ideal de objetividade total é inatingível e . . . os métodos clássicos, que posam de guardiões desse ideal, na verdade o violam a cada momento; virtualmente nenhum desses métodos pode ser aplicado sem um generosa ajuda de juízo pessoal e pressuposto arbitrário (Howson & Urbach 2006, p.9).[11]

Trabalhos recentes da estatística do erro não deixam essa questão sem resposta, dando continuidade ao debate.

Conclusão

A tensão entre bayesianismo e estatística do erro se expressa nos diferentes critérios apregoados por esses enfoques para realizar inferências. De ponto de vista bayesiano, o que importa é conhecer a probabilidade posterior da hipótese em exame. Em algumas situações, a probabilidade prévia será idêntica à frequência e a probabilidade posterior poderá sustentar uma inferência eventualmente coincidente com a inferência que realizaria um adepto da estatística do erro. Em outras situações, porém, quando se está diante de um caso de probabilidade prévia como expressão do grau de crença do pesquisador na hipótese, as inferências avalizadas pelo bayesianismo e pela estatística do erro podem levar a resultados completamente diferentes.

A depender dos graus de probabilidade prévia atribuídos, uma hipótese pode ser inferida pelo bayesianismo sem que tenha passado por um teste severo segundo a concepção da estatística do erro. Para Howson e Urbach, os resultados a que leva o bayesianismo são coerentes e são confirmados pelo senso comum. O fato de eles eventualmente autorizarem inferências incorretas se mostra como um problema menor, pois o acúmulo de indícios corrigirá isso no médio ou longo prazo.

A difusão dos métodos bayesianos por áreas tão diversas como a filosofia da ciência, a teoria da decisão, a estatística e a análise de risco, entre outras, constitui uma mostra de que, a despeito da opinião de seus detratores, o bayesianismo não pode ser descartado sem maiores problemas. É forçoso reconhecer, porém, que, como toda e qualquer escola filosófica, o bayesianismo também apresenta limites. De fato, a aplicação do Teorema de Bayes, por si só, não garante a realização de inferências que levem a resultados corretos. Para o bayesiano, a veracidade ou falsidade de uma hipótese será declarada à luz da acumulação de indícios. O bayesiano acredita que, no médio prazo, os indícios indicarão o caminho correto. Mas esse é um risco que um adepto da estatística do erro não está disposto a correr.

Para a estatística do erro, o objetivo primordial é realizar inferências tão avalizadas quanto o possível. Se não se pode ter certeza absoluta, deve-se, ao menos, controlar da melhor maneira possível todos os erros e incertezas envolvidas no caso. Segundo esse enfoque, o avanço da ciência e o raciocínio científico somente podem ser compreendidos em sua inteireza através da análise da atividade experimental. Nada interessa menos do que o grau de crença do pesquisador na hipótese em exame. Nada interessa mais do que investigar as margens de erro de determinado experimento e, a partir desse dado, assegurar que qualquer inferência será feita de forma tão segura quanto possível.

Há, assim, na proposta de Mayo, importantes elementos de continuidade da metodologia popperiana, mas também de novidade em relação ao trabalho de Popper. Ela mantém deste particularmente a ideia de que o mais importante na atividade científica é submeter as hipóteses a um teste rigoroso e evitar ao máximo o erro. Por outro lado, diferente do dedutivismo anti-indutivo de Popper, Mayo inclui a probabilidade no raciocínio científico e, assim, admite algum papel para a indução em sua proposta. Além disso, diferentemente de Popper, que admitia qualquer hipótese a princípio para ser testada, sem critérios prévios de seleção, a estatística do erro traz em si uma heurística, com elementos de um método para selecionar previamente hipóteses que mereçam mais atenção do cientista.

Se alguma conclusão pode ser extraída dos argumentos precedentes é a de que o debate entre bayesianismo e estatística do erro está longe de estar encerrado. A continuidade desse debate se dará menos em função da capacidade de um lado de desacreditar completamente o outro e mais em razão de que, na tentativa de compreender o raciocínio científico, ambas as propostas têm objetivos distintos e exibem considerável valor interpretativo, cada uma a seu modo. Talvez a atividade científica seja complexa demais para ser descrita com base em apenas uma interpretação do raciocínio científico. O fato de que diferentes formas podem levar a diferentes resultados é apenas outra maneira de reconhecer essa complexidade.

Para a filosofia da ciência, trata-se de uma continuação instigante — agora com a ajuda de sofisticados recursos de análise probabilísticos — do debate que constitui o que de mais interessante ela produziu em sua história como disciplina filosófica: o problema da natureza do raciocínio científico e o papel da experiência na avaliação de teorias científicas.

Referências

Corfield, D.; Williamson, J. (eds.) 2001. *Foundations of Bayesianism*. Dordrecht; Boston; London: Kluwer.

De Finetti, B. 1989. Probabilism. *Erkenntnis* **31**: 169–223.

Earman, J. 1992. *Bayes of Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: The MIT Press.

Gillies, D. 1991. Intersubjective Probability and Confirmation Theory. *British Journal for the Philosophy of Science 42*: 513–533.

———. 2000. *Philosophical Theories of Probability*. New York: Routledge.

Good, I. 1976. The Bayesian Influence, or how to Sweep Subjectivism under the Carpet. In: Harper; Hooker (eds.) 1976, *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science. v. II*, Dordrecht: Reidel.

Hempel, C. 1945. Studies in the Logic of Confirmation. *Mind* **volume 54** n. 213 (Janeiro de 1945): 1–26.

———. 1966. *Philosophy of Natural Science*. Englewood Cliffs: Prentice Hall.

Howson, C. 1997. Error Probabilities in Error. *Philosophy of Science* **64**(Suplemento): S185–S194.

———. 2003[2000]. *Hume's Problem: Induction and the Justification of Belief*. Oxford: Clarendon.

Howson, C.; Urbach, P. 2006[1989]. *Scientific Reasoning: The Bayesian Approach*. Chicago: Open Court.

Jeffreys, H. 1961. *Theory of Probability*. Oxford: Clarendon.

Kuhn, T. 1970[1962]. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Mayo, D. 1996. *Error and the Growth of Experimental Knowledge*. Chicago/London: The University of Chicago Press.

———. 1997a. Response to Laudan and Howson. *Philosophy of Science 64*: 323–333.

———. 1997b. Error Statistics and Learning from Error: Making a Virtue of Necessity. *Philosophy of Science* **64**(Suplemento): S195–S212.
———. 2010. *Error and Inference*. Cambridge University Press.
———. 2011. Error Statistics. In: *Handbook of the Philosophy of Science. Volume 7: Philosophy of Statistics*. Elsevier.
Pearson, E. S. 1950. On Questions Raised by the Combination of Tests Based on Discontinuous Distributions. *Biometrike* **37**: 383–398.
Popper, K. 1957. The Propensity Interpretation of the Calculus of Probability and the Quantum Theory. In: Köner, S. (ed.) 1957, *Observation and Interpretation*, Bristol: Butterworths Scientific Publications, p.65–70.
———. 1959. The Propensity Interpretation of Probability. *British Journal for the Philosophy of Science* **10**(37): 25–42.
———. 2002[1959]. *The Logic of Scientific Discovery*. London: Routledge.
Portugal, A. 2004. Probabilidade e Raciocínio Científico. *Episteme* **18**: 19–40.
Ramsey, F.P. 1931. Truth and Probability. In: R. B. Braithwaite (ed.) 1931, *The Foundations of Mathematics and other Logical Essays*. London: Kegan, Paul, Trench, Trubner & Co., New York: Harcourt, Brace and Company, p.156–98.
Salmon, W. 1998. Rationality and Objectivity in Science or Tom Kuhn meets Tom Bayes. In: J. A. Cover; M. Curd; C. Pincock (eds.) 1998, *Philosophy of Science: the central issues*. W.W. Norton Company, p.551-583 .
Swinburne, R. 2001. *Epistemic Justification*. Oxford: Oxford University Press.

AGNALDO CUOCO PORTUGAL
Universidade de Brasília
agnaldocp@unb.br

BRENO HERMANN
Universidade de Brasília
brhermann@yahoo.com

Resumo. O artigo apresenta os principais elementos de dois enfoques alternativos para o uso do cálculo de probabilidades na análise do raciocínio científico: o bayesianismo e a estatística do erro. O debate entre essas correntes é um dos mais relevantes da filosofia da ciência contemporânea e constitui uma continuação do conhecido debate entre Popper e Kuhn. São apresentadas explicações do bayesianismo que outros enfoques têm dificuldade em justificar, especialmente o paradoxo dos corvos. Não obstante seus pontos positivos, o enfoque bayesiano não está imune a críticas, como a de que não soluciona adequadamente o problema de Duhem. A estatística do erro, em particular, propõe a aplicação de métodos estatísticos e do cálculo de probabilidades para explicar o raciocínio científico de forma radicalmente distinta do bayesianismo. Com isso, o debate iniciado por Popper e Kuhn tem uma continuação em termos probabilísticos e está longe de concluído.

Palavras-chave: Deborah Mayo; Colin Howson; raciocínio científico; bayesianismo; estatística do erro.

Notas

[1] Para uma interessante aproximação da abordagem de Kuhn com a do bayesianismo, ver Salmon (1998).
[2] Há propostas bayesianas que não são subjetivistas em termos de teoria da probabilidade, como Swinburne (2001), por exemplo. No entanto, os subjetivistas são certamente majoritários entre os bayesianos (cf. Gillies, 2000).
[3] Para se ter uma ideia dos vários campos de aplicação do bayesianismo, vale a pena consultar Corfield, David e Williamson, Jon (eds), *Foundations of Bayesianism*. Dordrecht: Kluwer Academic Publishers (2001). Para uma abordagem bayesiana do problema da indução, ver Howson (2000).
[4] Uma modesta e aparentemente isolada tentativa de divulgação entre nós foi Portugal (2004).
[5] Determinados autores bayesianos que antecederam a Howson e Urbach, como Jaynes e Rosencrantz, acreditavam que as probabilidades prévias, apesar de exibirem certo grau de subjetividade, podiam e deviam ter sua atribuição de valores constrangida por fatores objetivos. Howson e Urbach, porém, consideram que qualquer tentativa de constranger a liberdade do cientista ao atribuir probabilidades prévias é ilusória. Critérios supostamente objetivos seriam, para eles, arbitrários e de difícil justificação.
[6] Os casos bem sucedidos do bayesianismo como teoria da confirmação apresentados aqui são apenas uma pequena seleção. Para uma apresentação mais completa, ver Howson & Urbach (especialmente 1993) e John Earman (1992).
[7] Popper chegou a formular um conceito de probabilidade como propensão (ver Popper 1957 e 1959), mas seu emprego é bastante limitado e não é incluído em sua teoria da justificação científica, que é estritamente dedutivista.
[8] Imagine-se uma série de predições dedutíveis de h e de h'. Uma vez que essas predições se realizem e, portanto, confirmem ambas as hipóteses, é possível calcular as probabilidades posteriores para ambas as teorias, de modo que $P\left(\frac{h}{e_1 \wedge e_2 \wedge \ldots \wedge e_n}\right) = \frac{P(h)}{P(e_1 \wedge e_2 \wedge \ldots \wedge e_n)}$ e $P\left(\frac{h'}{e_1 \wedge e_2 \wedge \ldots \wedge e_n}\right) = \frac{P(h')}{P(e_1 \wedge e_2 \wedge \ldots \wedge e_n)}$. Ao combinar essas duas equações e eliminar o denominador comum, tem-se $P\left(\frac{h}{e_1 \wedge e_2 \wedge \ldots \wedge e_n}\right) = \frac{P(h)}{P(h')} \times \frac{P(h')}{P(e_1 \wedge e_2 \wedge \ldots \wedge e_n)}$. Como o valor do último elemento à direita da equação não pode ser maior do que 1, isso significa que a probabilidade posterior de h nunca poderá ser maior que o fator $P(h)/P(h')$. Ver Howson e Urbach 2006, p.95.
[9] Trata-se de uma variação da concepção subjetiva de probabilidade: a de probabilidade intersubjetiva, desenvolvida, por exemplo, em Gillies (1991).
[10] Na mesma direção vai Good (1976).
[11] "... the ideal of total objectivity is unattainable and ... the classical methods, which pose as guardians of that ideal, actually violate it at every turn; virtually none of those methods can be applied without a generous helping of personal judgment and arbitrary assumption".