PROBABILISMO E BAYESIANISMO EM EPISTEMOLOGIA

André Luiz de Almeida Lisbôa Neiva*

Resumo:
Este artigo pretende oferecer um quadro geral das bases teóricas do probabilismo e do Bayesianismo em epistemologia. Os principais conceitos e definições do cálculo de probabilidades são fornecidos na seção 2; uma distinção entre tipos de probabilidade encontra-se na seção 3; o modelo de graus de crença como probabilidades subjetivas e a sua representação em termos de quocientes de aposta são os temas da seção 4; duas restrições sincrônicas de racionalidade probabilística são fornecidas na seção 5; restrições diacrônicas, condicionalização §1 e §2, e o teorema de Bayes são explicados na seção 6; a teoria Bayesiana de confirmação é o objeto central da seção 7; alguns dos principais problemas e objeções são discutidos na seção 8; por fim, as considerações finais e um apêndice contendo a demonstração de alguns teoremas básicos do aparato probabilístico.

Palavras-chave:
Condicionalização, Confirmação, Graus de Crença, Probabilidade, Teorema de Bayes.

Abstract:
This paper aims to offer an overview of the theoretical grounds of probabilism and Bayesianism in epistemology. The main concepts and definitions of the probability calculus *are provided in section 2; a distinction between some kinds of probability is given in section 3; the account of degrees of belief as subjective probabilities and their representation in terms of betting quotients are the topics of section 4; two constraints on probabilistic rationality are provided in section 5; the diachronic constraints, conditionalization* §1 *and* §2*, and the Bayes' Theorem are explained in section 6; the Bayesian confirmation theory is the main concern of the section 7; some of the major problems and objections are discussed in section 8; so the final remarks and an appendix containing the proof of certain basic theorems of the probability* apparatus *are available in the end of this paper.*

Keywords:
Conditionalization, Confirmation, Degrees of Belief, Probability, Bayes' Theorem.

*Mestrando em Filosofia pela Pontifícia Universidade Católica do Rio Grande do Sul; Bolsista CNPq. E-mail: al.neiva@gmail.com.

Preliminares

Tornou-se bastante comum nos últimos anos o uso de ferramentas formais de outras áreas em epistemologia analítica: da lógica, da teoria de probabilidade, da computação, da semântica, entre outras. Em particular, a introdução de ideias e conceitos da teoria de probabilidade e do Bayesianismo em epistemologia tem aumentado significativamente[1]. Uma motivação especial se deve ao fato do *calculus* probabilístico fornecer um aparato matemático acurado para lidar com importantes temas epistemológicos: graus de crença modelados sob uma interpretação de probabilidades subjetivas ou *credences*, estruturas formais de revisão e atualização dos graus de crença, racionalidade e coerência probabilísticas, os conceitos de confirmação absoluta e incremental, entre outros.

Mas o que são o probabilismo e o Bayesianismo? Probabilismo é a concepção que propõe, diferentemente do modelo *mainstream* de crença em epistemologia[2], que crenças de agentes doxásticos são uma questão de graus, modeladas em termos de probabilidades subjetivas ou *credences*. Graus de crença, contudo, precisam satisfazer as regras do maquinário probabilístico para que sejam racionais ou, em sentido mais estrito, coerentes probabilisticamente. Por sua vez, Bayesianismo é uma teoria que estabelece modos pelos quais graus de crença devem ser modificados através do tempo, a saber, em conformidade com regras de revisão e atualização: e.g., o princípio de condicionalização estrita em combinação com o teorema de Bayes. Nessa perspectiva, tais regras oferecem um esquema de como agentes racionais devem mudar seus graus de crença em resposta a novas evidências adquiridas.

Axiomas, Regras e Definições do Cálculo de Probabilidades

Assumindo que $Pr(\,\cdot\,)$ é uma função de probabilidade e F é um campo ou álgebra, tal que F é um conjunto finito de proposições fechado sob implicação verofuncional sobre um conjunto de possibilidades $\Omega \neq \emptyset$ (F é um conjunto de subconjuntos de Ω), temos os seguintes axiomas de probabilidade (axiomas de Kolmogorov[3]), onde

[1] Segue uma lista de alguns textos representativos sobre probabilismo e Bayesianismo em epistemologia formal: Bovens e Hartmann (2003); Joyce (1998 e 2004); Hájek (2009); Hájek e Hartmann (2010); Hartmann e Sprenger (2011). Bradley (2015) e Weisberg (2015, secs. 1 e 2) oferecem um excelente material introdutório com enfoque no uso do maquinário Bayesiano-probabilístico em epistemologia formal.

[2] O modelo de crença *simpliciter*, ou modelo de crença tudo-ou-nada, não considera crenças de agentes doxásticos em termos de gradações em um determinado intervalo. Um agente S crê ou não (ou suspende o juízo), em termos absolutos, numa certa proposição p em um dado instante t.

[3] A axiomatização de Kolmogorov (1956, cap. 1, p. 2) é comumente apresentada na linguagem da teoria dos conjuntos. Pode-se, no entanto, traçar algumas equivalências entre os operadores da teoria dos conjuntos e os de lógica proposicional: $P \wedge Q$ corresponde a $P \cap Q$, $P \vee Q$ corresponde a $P \cup Q$ e $\neg P$ corresponde a $\bar{P}$. As regras do *calculus* expostas aqui são comumente apresentadas em diversos manuais clássicos sobre o assunto (Howson e Urbach, 2006, cap. 2, p. 13-44; Hacking, 2001, cap. 6, p. 58-68;

$Pr : F \longrightarrow \Re$ — tal que $p \in F$ e $q \in F$ — *sse*[4]:

(1) $0 \leq Pr(p) \leq 1$;
(2) $Pr(\Omega) = 1$;
(3) Se p e q são mutuamente exclusivas[5] (ou incompatíveis), $Pr(p \vee q) = Pr(p) + Pr(q)$.

O axioma (1) é conhecido como não-negatividade, o (2) como certeza ou normalização e o (3) como aditividade finita[6]. $\langle \Omega, F, Pr \rangle$ é conhecido como espaço de probabilidade (*probability space*), sendo $Pr(\,\cdot\,)$ uma função que mapeia os números reais do intervalo $[0,1]$ sobre F. Em outras palavras, os valores de probabilidade atribuíveis a proposições situam-se em um intervalo, $[0,1]$, que inclui os números reais.

O axioma (2), tipicamente, assume uma formulação que determina o valor máximo a tautologias. Uma variação de (2) pode ser formulada:

(2') Se p é tautológica, $Pr(p) = 1$.

Algumas consequências lógicas podem ser derivadas dos axiomas acima. Dentre elas, a probabilidade de $p \vee \neg p$ tem o valor aferido igual a 1. $Pr(p \vee \neg p) = 1$, uma vez que p e $\neg p$ são mutuamente exclusivas (e logicamente contraditórias) e, por conseguinte, pelo axioma (3), $Pr(p \vee \neg p) = Pr(p) + Pr(\neg p)$; mas, por sua vez, $p \vee \neg p$ é uma tautologia e, pelo axioma (2'), $Pr(p \vee \neg p) = 1$. Outra é se $Pr(p \vee \neg p) = 1$ e $Pr(p \vee \neg p) = Pr(p) + Pr(\neg p)$, então $Pr(p) + Pr(\neg p) = 1$. Portanto, $Pr(\neg p) = 1 - Pr(p)$.

Ademais, existem casos nos quais duas ou mais proposições são compatíveis, ou seja, não formam um conjunto de proposições mutuamente exclusivas e a sua conjunção é logicamente consistente. Em tal circunstância, alegadamente, a probabilidade da disjunção é o resultado da soma das probabilidades dos disjunctos com a subtração da probabilidade da conjunção. Há um teorema conhecido como *overlap* para proposições não mutuamente exclusivas. Formalmente, ele pode ser colocado como se segue:

$$Pr(p \vee q) = Pr(p) + Pr(q) - Pr(p \wedge q)$$

O axioma (3) simplesmente fixa o valor 0 para a probabilidade da conjunção em situações em que duas ou mais proposições, os *conjunctos*[7] da proposição molecular,

Hájek, 2011, sec. 1; Earman, 1992, cap. 2, p. 33-62).

4 Abreviação da expressão 'se e somente se'.

5 Duas ou mais proposições são mutuamente exclusivas quando apenas uma delas pode ser verdadeira.

6 O axioma (3) pode ser reformulado para aditividade contável. Se F é um conjunto fechado sob infinitas disjunções contáveis — $p_1, p_2, p_3, \ldots, p_n$ formam um conjunto infinito de proposições mutuamente exclusivas —, então $Pr(p_1 \vee p_2 \vee p_3 \vee \ldots\, p_n) = Pr(p_1) + Pr(p_2) + Pr(p_3) + \ldots + Pr(p_n)$.

7 O termo 'conjuncto' está sendo usado em sentido técnico aqui. *Conjunctos* são, por assim dizer, propo-

são incompatíveis em conjunto. Por exemplo, a probabilidade de contradições lógicas é igual a 0: $Pr(p \wedge \neg p) = 0$. Pelo axioma (2'), depreende-se que $Pr[\neg(p \wedge \neg p)] = 1$, visto que $\neg(p \wedge \neg p)$ é uma tautologia. No entanto, $Pr[\neg(p \wedge \neg p)] = 1 - Pr(p \wedge \neg p)$ e, como $Pr[\neg(p \wedge \neg p)] = 1$, conclui-se que $Pr(p \wedge \neg p) = 0$.

Existem outras regras importantes do cálculo de probabilidades. Se as proposições p e q são logicamente equivalentes, $p \equiv q$, então os valores de probabilidade delas são iguais[8]. Por exemplo, a probabilidade da proposição *para todo x, se x é corvo, então x é preto* é igual à probabilidade da proposição *para todo x, se x não é preto, então x não é corvo*[9].

No caso em que uma proposição acarreta logicamente a outra, o valor de probabilidade da proposição acarretada não pode ser menor do que o da proposição que a acarreta: se $p \vDash q$, então $Pr(p) \leq Pr(q)$. A única possibilidade, que está disponível no apêndice, de $Pr(q) = Pr(p)$ é sob a condição de que $Pr(\neg p \wedge q) = 0$.

Até aqui falamos apenas de probabilidade categórica, a saber, probabilidade incondicionada. É de bom alvitre, no entanto, introduzir a noção de probabilidade condicional, expressa na notação $Pr(\cdot \mid \cdot)$.

Suponha um dado de seis faces (lados) e com um mecanismo justo de jogadas (tentativas) — o conceito de justo (*fair*) significa que o mecanismo de jogadas não é viciado (não-enviesado) e os eventos em instâncias individuais são independentes entre si, isto é, os resultados em ocorrências $t_1, t_2, \ldots, t_n$ não se influenciam entre si —, a probabilidade de sair/obter o número 2 em uma determinada instância particular é $\frac{1}{6}$; nesse exemplo, $Pr(obter\ o\ resultado\ 2) = \frac{1}{6}$, uma vez que são seis resultados possíveis, a saber, o conjunto de possibilidades encerra 6 resultados simples ou o espaço de amostra (*sample space*) é $\Omega = \{1,2,3,4,5,6\}$. Agora, usando o mesmo dado, considere a probabilidade de sair/obter o número 2 condicionado ao fato de sair/obter um número par, nomeadamente, $Pr(obter\ 2 \mid obter\ par)$. O resultado de sair um número par encerra três possibilidades de resultados, o subconjunto $\{2,4,6\}$ do conjunto maior Ω, o que lhe confere a probabilidade de $\frac{1}{2}$, já que ou é par ou é ímpar. Dessa maneira, $Pr(obter\ 2 \mid obter\ par) = \frac{1}{3}$ porque dos três resultados possíveis de números pares somente um resultado é o número 2; por assim dizer, somente o número 2 satisfaz as duas condições: ser par e ser propriamente o número 2.

A definição de probabilidade condicional, $Pr(p \mid q)$ significa a probabilidade de p condicionada em q (ou a probabilidade de p, dado que q), pode ser apresentada

sições atômicas que conectadas pelo operador de conjunção ($\wedge$) formam uma proposição molecular ou composta.

8 A demonstração de tal teorema está disponível no apêndice deste texto.

9 $[\forall x(Rx \rightarrow Bx)] \equiv [\forall x(\neg Bx \rightarrow \neg Rx)]$ é uma equivalência lógica conhecida como transposição. Por conseguinte, $Pr[\forall x(Rx \rightarrow Bx)] = Pr[\forall x(\neg Bx \rightarrow \neg Rx)]$.

formalmente. Para quaisquer proposições p e q e assumindo que $Pr(q) > 0$:

$$Pr(p \mid q) = \frac{Pr(p \wedge q)}{Pr(q)}$$

Outra definição fundamental do cálculo de probabilidades é a de independência. Duas proposições p e q são independentes, assumindo que $Pr(p) > 0$ e $Pr(q) > 0$, se e somente se $Pr(p \mid q) = Pr(p)$.

Um corolário das definições de probabilidade condicional e independência é a que se segue. Suponha que p e q são independentes entre si e que $Pr(p) > 0$ e $Pr(q) > 0$. Assim, $Pr(p \mid q) = Pr(p)$. Mas, pela definição de probabilidade condicional, $Pr(p \mid q) = \frac{Pr(p \wedge q)}{Pr(q)}$. Por conseguinte, $Pr(p) = \frac{Pr(p \wedge q)}{Pr(q)}$. Portanto, $Pr(p \wedge q) = Pr(p) \times Pr(q)$.

Além disso, podemos dizer que se p e q são independentes, então q é irrelevante para a probabilidade de p e, igualmente, p é irrelevante para a probabilidade de q. Em outras palavras, q não contribui no valor de probabilidade de p e vice-versa. Trata-se de uma relação simétrica. Se $Pr(p \mid q) = Pr(p)$, então $Pr(q \mid p) = Pr(q)$ e se $Pr(q \mid p) = Pr(q)$, então $Pr(p \mid q) = Pr(p)$. Portanto, $Pr(p \mid q) = Pr(p)$ *sse* $Pr(q \mid p) = Pr(q)$.

Podemos, nesta altura, falar sobre a noção de probabilidade total. Essa e a definição de probabilidade condicional são decisivas para o Bayesianismo. A rigor, consegue-se a demonstração do teorema de Bayes pela combinação das duas. Para quaisquer proposições p e q e assumindo que $1 > Pr(q) > 0$:

$$Pr(p) = Pr(q) \times Pr(p \mid q) + Pr(\neg q) \times Pr(p \mid \neg q)$$

Finalmente, precisamos considerar que nunca a probabilidade da conjunção pode ser maior do que a probabilidade de qualquer um dos seus *conjunctos* isolados, a saber, $Pr(p \wedge q) \leq Pr(p)$ e $Pr(p \wedge q) \leq Pr(q)$. Todavia, em circunstâncias em que p acarreta logicamente q, a probabilidade de p é igual à probabilidade da conjunção $p \wedge q$. Formalmente, se $p \vDash q$, então $p \equiv (p \wedge q)$ e, logo, $Pr(p) = Pr(p \wedge q)$.

Tipos de Probabilidade

Apresentamos na seção anterior o maquinário formal probabilístico básico, com exceção do teorema de Bayes e das regras de condicionalização (objetos da seção 6). Nesta seção, vamos nos ocupar com três conceitos ou tipos distintos de probabilidade. Assim, de acordo com a classificação proposta por D. H. Mellor (2005, cap. 1, p. 8-13), podemos fazer uma distinção entre probabilidade física ou natural (*chance*), probabilidade epistêmica e probabilidade subjetiva ou pessoal (*credences*)[10].

Considere as proposições abaixo:

(a) Este dado é justo, a probabilidade de sair o resultado de número 2 é $\frac{1}{6}$;
(b) É muito provável que o Universo teve um início;
(c) Estou confiante, provavelmente o Inter vencerá o Gre-nal, aposto $4:1$ a favor do Inter.

Com efeito, (a) refere-se a um tipo de probabilidade conhecida como *chance* ou probabilidade física. Presumivelmente, tal tipo de probabilidade descreve uma propriedade física do mundo. É, nesse sentido, objetiva e independente do que se passa na vida mental de um agente, isto é, independente se alguém em algum momento a considerou e soube ou sabe a respeito dela. Há diferentes explicações sobre o que são probabilidades físicas. Uma abordagem baseada em *frequência relativa* interpreta (a) da seguinte maneira: em uma longa sequência de um total de *n* jogadas repetidas de um determinado dado — desde que *n* satisfaça um limiar de valor (*threshold*) que caracterize uma longa sequência de jogadas —, a probabilidade de sair o resultado de número 2 é a *frequência relativa* desse resultado no conjunto de *n* jogadas. É a proporção que tal resultado acontece no conjunto total de jogadas em uma longa sequência finita. Uma explicação discrepante, todavia, recorre ao uso do conceito de *propensão* ou *disposição*. Nessa perspectiva, (a) é interpretada como expressando uma determinada tendência de sair o resultado de número 2 frente às outras possibilidades. Supondo que a probabilidade de cada um dos possíveis resultados, $\{1,2,3,4,5,6\}$, é $\frac{1}{6}$, haveria uma distribuição equânime de probabilidade para cada um dos resultados. Isso envolveria considerações sobre a física do dado: o *modo* como ele é jogado, a *simetria* das disposições para cada resultado possível. Podemos testar *empiricamente* um mecanismo de jogadas (do dado) a fim de verificarmos se o resultado de 2 ocorre em proporção igual ou não aos outros resultados. De todo modo, (a) descreve como o mundo funciona e é verdadeira ou falsa independentemente se podemos saber ou não de tal fato[11].

Por seu turno, (b) não descreve uma característica física do mundo. Ela enuncia, alegadamente, a probabilidade de uma hipótese. Por exemplo, a hipótese cosmológica do *Big Bang* é uma explicação amplamente aceita na comunidade científica sobre o surgimento do Universo (em disputa com a hipótese do estado estacionário). Hipóteses científicas, por sua vez, estão ancoradas em evidências e, admitidamente, pode-se mensu-

[10] Existem outras tipologias na literatura sobre probabilidade. Carnap (1962, cap. 2, p. 19) diferencia somente dois tipos de probabilidade: $tipo_1$ é o grau de confirmação de uma hipótese, dado um conjunto de evidências que a suporta e o $tipo_2$ é entendido como *frequency-type*, a saber, a frequência relativa na qual um evento do mundo físico ocorre.

[11] Sobre interpretações e análises da função probabilística em geral, ver Hájek (2011). Gillies (2000) oferece um bom material acerca de interpretações frequentistas (cap. 5) e propensitivas (caps. 6 e 7).

rar *quão bem*, e *em que grau*, evidências disponíveis suportam diferentes hipóteses. Na verdade, (b) precisa ser reformulada, pois parece correto afirmar que, considerando um determinado conjunto de evidências, a hipótese em questão em (b) é provável. Está implícito essa relação de suporte. Por isso, é mais apropriado que:

(b') Dado todo o conjunto disponível de evidências astronômicas, é muito provável que o Universo teve um início.

Destarte, (b') reporta-se a um tipo diferente de probabilidade: epistêmica ou evidencial. Probabilidades epistêmicas estão condicionadas em evidência(s), o que não é necessariamente verdadeiro para as probabilidades física e subjetiva. Essa é uma marca distintiva da natureza da probabilidade epistêmica. Como veremos mais adiante, graus de confirmação interpretados como probabilidades epistêmicas, que pretendem medir *absolutamente* ou *incrementalmente* o suporte que um conjunto de evidências oferece a uma hipótese, estão condicionados em toda evidência disponível: $e_t = e_1 \wedge e_2 \wedge \ldots e_n \wedge k$, a saber, partes da evidência + conhecimento de fundo (*background knowledge*). Nessa perspectiva, tais probabilidades mensuram *quão forte* um conjunto total de evidências disponíveis, e_t, torna uma determinada hipótese, h, provável, isto é, em que *grau* a confirma, ou a desconfirma, ou, ainda, se a evidência é neutra em relação à hipótese[12].

Por último, (c) descreve um tipo distinto: probabilidades subjetivas ou *credences*. A afirmação (c) não descreve, em absoluto, uma característica do mundo e pode ou não estar apoiada em evidências. De todo modo, (c) expõe uma disposição de um agente em apostar num certo grau que uma certa proposição é verdadeira. É uma concepção intimamente vinculada à teoria da decisão e tais probabilidades estão associadas a contextos que comportamentos de apostas (*betting behaviour*) ocorrem. Nesse sentido, quocientes de aposta (*betting quotients*) oferecem uma representação numérica a graus de crença. Mostraremos na próxima seção como é possível equipará-los. Em linhas gerais, probabilidades subjetivas determinam em que medida e *quão fortemente* um agente doxástico crê em proposições. A alegação de fundo é que cremos mais ou menos fortemente em proposições e que podemos modelar probabilisticamente tais atitudes doxásticas. Pode-se representar que a crença em uma proposição p tem um determinado valor ε dentro de um intervalo numérico, $[0,1]$, relativamente a um agente S num dado instante t.

[12] É matéria de controvérsia se probabilidades epistêmicas podem ser *lógicas*. Se isso é o caso, então graus de confirmação poderiam ser estabelecidos por uma lógica indutiva, de modo análogo às relações de acarretamento (*entailment*) da lógica dedutiva, funcionando como uma extensão das lógicas proposicional e de primeira-ordem.

Graus de Crença e Quocientes de Aposta

Uma das principais teses do probabilismo e do Bayesianismo é a de que existem graus de crença e que é possível representá-los formalmente sob a estrutura do *calculus* probabilístico, *i.e.*, graus modelados em termos de *credences*. Assim, podemos ter uma variedade de graus de crença em relação a diferentes proposições e, *ipso facto*, podemos ter maior grau de crença em uma determinada proposição do que em comparação aos graus de crença que temos em outras proposições.

O modelo de graus de crença, em todo caso, está conectado a disposições e comportamentos de aposta: a inclinação de agentes apostarem sobre a verdade de proposições[13]. Quocientes de aposta (*betting quotients*) oferecem uma medida numérica à concepção de graus de crença. O grau ε de crença na proposição r de um agente S corresponde à disposição de S apostar que r é verdadeira ou falsa. Por exemplo, se alguém tem um quociente de aposta de $\frac{3}{5}$ (ou .6) que *vai chover hoje à noite em Porto Alegre* (r) — e de $\frac{2}{5}$ (ou .4) que *não vai chover hoje à noite em Porto Alegre* ($\neg r$) —, significa dizer que tal agente tem um grau de crença maior de que *vai chover hoje à noite em Porto Alegre* do que não[14].

Vamos propor um experimento mental. Suponha, por exemplo, que dois agentes S_1 e S_2 façam uma aposta. S_1 aposta N que r e S_2 aposta M que $\neg r$. O montante (ou totalização) das apostas de S_1 e S_2, $N+M$, é chamado de *stake*. A fórmula do quociente de aposta de cada um dos apostadores envolvidos, por sua vez, é a *razão* (*ratio*) do valor de aposta pelo *stake*:

$$(i)\ Quociente\ de\ aposta\ de\ S_1 = \frac{N}{(N\ +\ M)}$$

$$(ii)\ Quociente\ de\ aposta\ de\ S_2 = \frac{M}{(N\ +\ M)}$$

Supondo a grandeza u de *utiles*[15], se $N = 4u$ e $M = 1u$, então S_1 aposta $4u$ a favor de r e S_2 aposta $1u$ a favor de $\neg r$, o que perfaz um total apostado (*stake*) de $5u$. Por conseguinte, se r é o caso, então S_1 tem um lucro de $1u$. Se $\neg r$ é o caso, então S_2 tem um lucro de $4u$. A perda (monetária ou de utilidade) de S_2 é de $1u$ se r é o caso e a de S_1 é de

[13] Tradição de interpretação subjetiva da função de probabilidade que remete a Frank P. Ramsey em *Truth and Probability* (1950, [1926]) e Bruno de Finetti em *Foresight: its Logical Laws, its Subjective Sources* (1964, [1937]).

[14] Em tal situação, o agente é, sob certo aspecto, *coerente probabilisticamente* em uma dada instância de tempo t, pois $Pr(r \vee \neg r) = 1$. Falaremos de condições sincrônicas de coerência probabilística no próximo tópico.

[15] *Utiles* é uma medida artificial de utilidade. Podemos, igualmente, falar de uma moeda como o valor implicado na aposta, o real R$, por exemplo. Uma proposta atraente é a conversão de vários tipos diferentes de utilidade em uma mesma grandeza.

$4u$ se $\neg r$ é o caso. Portanto, o quociente de aposta de S_1 é de $\frac{4}{5}$ a favor de r e o de S_2 é de $\frac{1}{5}$ a favor de $\neg r$.

Algumas considerações são decisivas. Primeira, apostar a favor de r equivale a apostar contra $\neg r$ e apostar a favor de $\neg r$ equivale a apostar contra r. No exemplo acima, S_1 aposta $\frac{4}{5}$ contra $\neg r$ e S_2 aposta $\frac{1}{5}$ contra r. Segunda, apostas envolvem risco: o apostador pode ganhar ou perder. No entanto, não significa que estamos falando de apostadores atuais e comportamentos concretos de aposta. O benefício da conversa sobre quocientes de aposta consiste exclusivamente na obtenção de uma representação numérica para graus de crença. Na verdade, as apostas envolvidas no experimento mental anterior são meramente *hipotéticas*. Terceira, podemos converter quocientes de aposta, que representam graus de crença, para *odds*, talvez mais popular entre jogadores e apostadores do mundo anglo-saxônico. Assim, ao passo que S_1 aposta $4:1$ a favor de r, S_2 aposta $4:1$ contra r. Quarta, está pressuposto no experimento mental que os quocientes de aposta são justos (*fair betting quotients*). Significa dizer que um agente S considera que não há vantagem (ou desvantagem) nem em apostar a favor de r com quociente de ε, nem em apostar contra r com quociente de $1-\varepsilon$; isto vale para ambos agentes, S_1 e S_2, do experimento acima[16]. Quinta, é um *desideratum* epistêmico que graus de crença, representados por probabilidades subjetivas, estejam condicionados em (boa) evidência, o que lhes confere, em acréscimo de outras restrições, uma chancela de *graus racionais de crença*. No entanto, há disputa, como veremos, se os *priors*, graus de crença anteriores, podem ou não estar condicionados em alguma restrição adicional, *e.g.* o princípio de indiferença, como defende o Bayesianismo objetivo. Sexta, o grau de crença χ de um agente S pode ter um valor preciso, *a sharp value*, $\chi \in [0,1]$ (ou $0 \leq \chi \leq 1$, por exemplo $\chi = .5$) , ou impreciso, *an unsharp value*, onde χ é todo intervalo $[a,b]$ tal que $[a,b] \subseteq [0,1]$. Neste último caso, um intervalo inteiro, um subconjunto de um intervalo maior, é tomado como sendo o grau de crença de um agente.

Restrições Sincrônicas

Falamos na seção anterior como é possível atribuir uma representação quantitativa a graus de crença, i.e., pelos quocientes de aposta. Precisamos, todavia, perguntarnos por que razão, e sob quais condições, graus de crença (e quocientes de aposta) precisam ser probabilisticamente coerentes: por que graus de crença devem obedecer o *calcu-*

[16] Pode-se conceber quocientes de aposta condicionais: apostar em r, dado que s ou apostar em $\neg r$, dado que s. São cenários similares ao experimento descrito acima, porém com a condição de que s seja o caso. Caso s não seja o caso, a aposta é, por assim dizer, anulada: os jogadores envolvidos nem ganham e nem perdem nada.

lus probabilístico?[17]

Suponha, agora, que um agente *S* tenha os seguintes quocientes de aposta: $\frac{3}{5}$ a favor de *r* (.6 a favor de *r*) e $\frac{3}{5}$ contra *r* (.6 contra *r*). Com base nos quocientes de *S*, um dono de uma casa de apostas muito matreiro oferece um contrato a *S* e determina sob que condições ele pode apostar: *S* aposta R$ 3 a favor de *r* (contra $\neg r$) podendo lucrar R$ 2 e *S* aposta R$ 3 contra *r* (a favor de $\neg r$) podendo lucrar R$ 2 igualmente. Se *S* aceitar a proposta oferecida, ele *certamente* perderá, de acordo com a tabela abaixo na coluna do *Saldo*:

	Aposta a favor de r	*Aposta contra r*	*Saldo*
r	+ R$ 2	− R$ 3	− R$ 1
$\neg r$	− R$ 3	+ R$ 2	− R$ 1

O cenário acima mostra que um conjunto de quocientes de aposta incoerente probabilisticamente — ou seja, que viola o maquinário probabilístico — está suscetível a um contrato de perda garantida ou a um *Dutch Book*. Analogamente, de acordo com o argumento do *Dutch Book*, um conjunto de graus de crença (de um agente doxástico) que não satisfaz o *calculus* de probabilidade é, da mesma forma, incoerente. Em outras palavras, se os graus de crença de um agente doxástico *S* não obedecem os axiomas e regras de probabilidade, então *S* está vulnerável a um contrato de perda garantida e poder-se-ia alegar incoerência no seu conjunto de graus de crença[18].

O argumento do *Dutch Book* é frequentemente construído como se segue[19]. Um conjunto de graus de crença *D* de um agente *S* deve corresponder a um conjunto de quocientes de aposta *Bq* de *S*. Se *Bq* de *S* viola o *calculus* de probabilidade, tornando *Bq* incoerente probabilisticamente, então há um *Dutch Book* envolvendo tais quocientes contra *S*. Por seu turno, se há um *Dutch Book* contra *Bq* de *S*, então *Bq* de *S* está vulnerável a um contrato que resulta em perda garantida. Se *Bq* de *S* está vulnerável a um contrato que resulta em perda garantida, então *S* é irracional. Portanto, supondo a correspondência

[17] O argumento do *Dutch Book* é considerado o argumento clássico a favor do probabilismo. Existem, entretanto, argumentos com estratégias diferentes a favor de tal tese. Joyce (1998) oferece um argumento baseado na acurácia dos graus de crença, onde a acurácia é medida, para qualquer proposição p, pela diferença entre uma função de valor de verdade $w(p)$ e uma função de graus de crença $Cr(p)$. Isso nos fornece um *score* e quanto maior o *score*, menos acurado é o grau de crença. Maher (1993) articula um argumento baseado no teorema representacional e na noção de preferências, onde estas precisam obedecer certas restrições básicas de racionalidade. Para um *survey* de tais posições, recomenda-se ver Hájek (2009) e Bradley (2015, cap. 3).

[18] É o que se entende por *Dutch Book Theorem*. Mas se os graus de crença (e os quocientes de aposta) de um agente *S* satisfazem o *calculus* probabilístico, então não existe um conjunto de apostas que torna *S* vulnerável a um contrato de perda garantida. Isso é conhecido como *Converse Dutch Book Theorem*, mais informações em Vineberg (2011, sec. 1.3) e Hájek (2009, p. 233). Provas do *Dutch Book Theorem* em Earman (1992, p. 39-40) e do *Converse Dutch Book Theorem* em Kemeny (1955, p. 268-269).

[19] Versões similares ao argumento apresentado aqui em Bradley (2015, cap. 3, p. 32), Hacking (2001, p. 165) e Huber (2009, p. 5-6).

entre Bq e D, conclui-se por silogismo hipotético que se o conjunto de graus de crença D de S viola o *calculus* de probabilidade, tornando D incoerente probabilisticamente, então S é irracional.

Esse argumento, na sua versão *default*[20], determina se um agente é ou não coerente, no que diz respeito aos seus graus de crença, em uma determinada instância particular de tempo. Como sugerem Hájek e Hartmann (2010, p. 95), trata-se de uma dimensão *sincrônica* de racionalidade e coerência probabilística. A propósito, os axiomas e regras explicados na seção 2 também referem-se a uma dimensão *sincrônica*: grau ε de crença, modelado sob o *calculus* probabilístico, de um agente S em uma instância de tempo t; e não como tal grau se atualiza através do tempo. No exemplo acima, S tem graus de crença de $Pr(r) = .6$ $(\frac{3}{5})$ e $Pr(\neg r) = .6$ $(\frac{3}{5})$ em t. S é incoerente em t, pois a probabilidade de $r \vee \neg r$ não pode superar o grau máximo de 1, de acordo com o primeiro axioma de probabilidade. Em sentido rigoroso, nos termos dos quocientes de S e do cenário construído, S tem um grau de 1.2 na probabilidade de $r \vee \neg r$, uma vez que r e $\neg r$ são mutuamente exclusivas e, pelo axioma de aditividade finita, $Pr(r \vee \neg r) = Pr(r) + Pr(\neg r) = .6 + .6 = 1.2$. Entretanto, além do fato do grau de 1.2 superar o valor *maximum* do intervalo $[0,1]$, o que significa infringir o primeiro axioma, S viola o axioma sobre o valor atribuído a tautologias[21]. Como $r \vee \neg r$ é uma tautologia ou verdade lógica, S teria que lhe conferir o valor preciso de 1.

Uma segunda restrição sobre graus de crença remete a um princípio formulado por David Lewis (1980, p. 266-267), o qual é conhecido como *princípio principal*. Tal princípio estabelece que se S tem boas evidências, e nenhuma evidência contrária, sobre a probabilidade física de um evento, veiculada por uma proposição p, com valor específico de ε, então o grau de crença de S de que p precisa se ajustar ao valor de ε. Assim, o princípio principal determina que graus de crença, *credences*, precisam corresponder a probabilidades físicas, quando estas estão disponíveis como evidência ao agente, num certo valor ε que indica a probabilidade de sua ocorrência no mundo. Formalmente, o princípio principal apresenta-se nos seguintes termos: $Cr(p \mid Ch(p) = \varepsilon) = \varepsilon$, onde $Cr(\ \cdot\ )$ é um função de probabilidade subjetiva (*credence*) e $Ch(\ \cdot\ )$ é uma função de probabilidade física (*chance*).

Considere que S tem boas evidências, supondo a ausência de qualquer evidên-

[20] Há uma literatura ampla sobre o assunto (para citar alguns, Gillies, 2000, cap. 4, p. 58-64; Jeffrey, 1983, cap. 4, p. 60-62). Também existem vários outros tipos de argumentos do *Dutch Book*. Vineberg (2011) discute em detalhes versões alternativas do *Dutch Book*: versões diacrônicas (com intuito de justificar princípios de condicionalização), versão para aditividade contável, entre outras.

[21] As violações exploradas são, por assim dizer, duas violações do *calculus* probabilístico básico, ou seja, dos dois primeiros axiomas de Kolmogorov. Cenários de violação do axioma de aditividade finita são mais engenhosos. Como tomaria muito espaço construí-los aqui, sugere-se ver Bradley (2015, cap. 3, p. 31-32), Hacking (2001, cap. 14, p. 166-167) e Vineberg (2011, sec. 1.1).

cia contrária, de que a frequência relativa em uma longa sequência de n jogadas finitas, relativas a um determinado mecanismo de jogadas, é distribuída igualmente entre as duas possibilidades de resultado, isto é, $\Omega = \{cara, coroa\}$. Por conseguinte, se p_1 enuncia a possibilidade de resultado de sair *cara* e p_2 enuncia a do resultado de sair *coroa*, temos $Ch(p_1) = .5$ e $Ch(p_2) = .5$. Os graus de crença de S, portanto, devem concordar com as probabilidades físicas, que estão disponíveis como evidência a S. Temos, neste caso, $Cr(p_1 \mid Ch(p_1) = .5) = .5$ e $Cr(p_2 \mid Ch(p_2) = .5) = .5$.

Restrições Diacrônicas e Teorema de Bayes

Ao passo que o contrato de perda garantida, ou *Dutch Book*, e o princípio principal são condições *sincrônicas* de racionalidade probabilística, elementos fulcrais do probabilismo, regras de atualização ou revisão dos graus de crença, propostas pelo Bayesianismo, são condições *diacrônicas* de racionalidade probabilística, como bem assinalam Hájek e Hartmann (2010, p. 95) e Huber (2009, p. 7). Tais modelos ou regras de revisão estabelecem, fundamentalmente, processos pelos quais um agente doxástico S atualiza o seu grau χ de crença em uma certa proposição h de uma instância de tempo t para uma instância de tempo t' quando alguma nova evidência ou informação é obtida por S.

Condicionalização estrita ou simples, doravante *regra de condicionalização* §1, é um princípio de revisão que trata da relação dos graus de crença em conjunto através do tempo, especificamente quando um agente S torna-se certo, com grau máximo 1, de uma determinada proposição evidencial e, sendo que uma proposição h está condicionada em e, isto é, $Pr_t(h \mid e)$. Assim, S deve atualizar o seu grau de crença posterior em h, a saber, $Pr_{t'}(h)$. Numa formulação mais precisa, a *regra de condicionalização* §1 pode ser apresentada da seguinte maneira: para quaisquer proposições h e e, se um agente S com graus de crença iniciais $1 > Pr_t(e) > 0$, $Pr_t(h \mid e) = \chi$ e $1 > Pr_t(h) > 0$ recebe evidência de tal modo que torna-se certo de e, então S deve igualar seu grau de crença inicial em h condicional em e, $Pr_t(h \mid e)$, ao seu grau de crença posterior em h, a saber, $Pr_t(h \mid e) = \chi = Pr_{t'}(h)$.

Nem sempre, no entanto, obtemos informações ou evidências que tornam nossos graus de crença tão certos, isto é, não necessariamente nos tornamos certos de e, $Pr_{t'}(e)$, em diversas situações. Por isso, Richard Jeffrey (1983, p. 172) propôs um princípio de condicionalização para revisão dos graus de crença com valores inferiores à certeza, ou seja, $Pr_{t'}(e)$ com valor não-maximal. Pode-se, desse modo, formular a *regra de condicionalização* §2: para quaisquer proposições h e e, se um agente S com graus de crença iniciais $1 > Pr_t(e) > 0$ e $1 > Pr_t(h) > 0$ recebe evidência de tal modo que altera o seu grau inicial em e para o valor χ, $Pr_{t'}(e) = \chi$, então S deve revisar o seu grau de crença inicial em h, $Pr_t(h)$, para um grau posterior tal que $Pr_{t'}(h) = [\chi \times Pr_t(h \mid$

$e)] + [(1-\chi) \times Pr_t(h \mid \neg e)]$.

A equação contida no consequente da *condicionalização* §2 converte-se na fórmula $Pr_{t'}(h) = [Pr_{t'}(e) \times Pr_t(h \mid e)] + [Pr_{t'}(\neg e) \times Pr_t(h \mid \neg e)]$. Ademais, a *regra de condicionalização* §2 é redutível, plausivelmente, na *regra de condicionalização* §1 sob uma determinada condição. Assim, considerando que $\chi = 1$, $Pr_{t'}(e) = 1$, temos $Pr_{t'}(h) = [1 \times Pr_t(h \mid e)] + [0 \times Pr_t(h \mid \neg e)]$ e, portanto, $Pr_{t'}(h) = Pr_t(h \mid e)$.

Embora os princípios de *condicionalização* §1 e §2 sejam peças essenciais, ainda nos falta introduzir um teorema central demonstrado pelo matemático e clérigo Thomas Bayes[22] (c. 1701-1761). A despeito de algumas diferenças entre o teorema original e a sua formulação mais contemporânea, ele é tomado aqui como uma ferramenta para modelar graus subjetivos de crença de agentes racionais.

Dada uma partição[23] de duas hipóteses h e $\neg h$ competidoras e desde que $Pr(e) > 0$, enuncia-se que

$$\langle Bayes\ §1 \rangle\ Pr(h \mid e) = \frac{Pr(h) \times Pr(e \mid h)}{[Pr(h) \times Pr(e \mid h)] + [Pr(\neg h) \times Pr(e \mid \neg h)]}$$

sendo $Pr(h)$ e $Pr(\neg h)$ as probabilidades anteriores, os *priors*, que representam os graus inicias de crença de um agente em h e $\neg h$, $Pr(e \mid h)$ o *likelihood*, o qual representa o grau de crença em e condicionado em h, e, finalmente, $Pr(h \mid e)$ a probabilidade posterior como o grau de crença do agente em h condicionado em e após realizar probabilisticamente o raciocínio.

O teorema *Bayes* §1 pode ser generalizado para uma partição finita de hipóteses competidoras $\{h_1, h_2, \ldots, h_n\}$, dado que $i = 1$, $i = 2$, ..., $i = n$. Assim, para toda h_i na partição, admitindo que $Pr(e) > 0$ e $Pr(h_i) > 0$, segue-se que

$$\langle Bayes\ §2 \rangle\ Pr(h_i \mid e) = \frac{Pr(h_i) \times Pr(e \mid h_i)}{\sum_j [Pr(h_j) \times Pr(e \mid h_j)]}$$

sendo o denominador composto pela soma dos produtos das probabilidades anteriores com os *likelihoods* de cada hipótese da partição.

Vamos propor um exemplo[24] de atualização que combina *condicionalização* §1 e *Bayes* §1. Supondo que um agente S tem algumas informações iniciais sobre uma

[22] A demonstração de tal teorema foi publicada apenas postumamente por Richard Price, amigo de Thomas Bayes, no ano de 1763. O artigo original de Bayes (2002 [1763], p. 117-149), *An Essay towards solving a Problem in the Doctrine of Chances*, está disponível em um volume editado por Richard Swinburne, juntamente com outros textos de acesso ao Bayesianismo. Uma ótima avaliação da obra de Bayes está disponível em Earman (1992, cap. 1, p. 7-31).

[23] Conjunto de proposições mutuamente exclusivas, apenas uma pode ser verdadeira, e reunidamente exaustivas, ao menos uma é verdadeira.

[24] Exemplo similar ao *Harvard Medical School Test*, disponível em Howson e Urbach (2006, cap. 2, p. 22-26).

determinada doença que acomete a população em geral: 1 a cada 2.000 indivíduos, isto é, $Pr_t(h) = 0.0005$. S tem outras informações que dizem respeito à confiabilidade do exame que atesta positivo ou negativo sobre a doença em questão. Dado que alguém tem a doença, a probabilidade de resultado negativo no teste é zero, $Pr_t(\neg e \mid h) = 0$, e dado que alguém não tem a doença, a probabilidade de resultado positivo no teste é de 0.15, ou seja, $Pr_t(e \mid \neg h) = 0.15$. S realiza o teste, recebe o resultado de positivo e torna-se certo de e, ou seja, $Pr_{t'}(e) = 1$. Ademais, se $Pr_t(h) = 0.0005$, então $Pr_t(\neg h) = 0.9995$; se $Pr_t(\neg e \mid h) = 0$, então $Pr_t(e \mid h) = 1$, pois $Pr_t(\neg e \mid h) = 1 - Pr_t(e \mid h)$. S deve computar as informações, dado que o teste é positivo, e atualizar o seu grau anterior a um novo grau posterior. Combinando *condicionalização* §1 e *Bayes* §1, segue-se que

$$Pr_{t'}(h) = Pr_t(h \mid e) = \frac{Pr_t(h) \times Pr_t(e \mid h)}{[Pr_t(h) \times Pr_t(e \mid h)] + [Pr_t(\neg h) \times Pr_t(e \mid \neg h)]}$$

$$Pr_{t'}(h) = Pr_t(h \mid e) = \frac{0.0005 \times 1}{[0.0005 \times 1] + [0.9995 \times 0.15]} \approx 0.0033$$

o que confere uma baixa probabilidade de S ter a doença. Em outros termos, se S atualiza de modo correto os seus graus, sendo sensível à obtenção de nova evidência e como prescrevem as regras de *condicionalização* §1 e *Bayes* §1, então S tem um grau posterior muito baixo sobre a hipótese de estar com a doença em apreço. Neste caso, como o agente realiza corretamente a atualização, pode-se afirmar que ele é irrepreensível do ponto de vista diacrônico-probabilístico.

Confirmação

A teoria Bayesiana de confirmação defende que a relação de suporte entre uma proposição evidencial, ou conjunto de evidências, e uma hipótese é medida probabilisticamente, ou seja, *graus de confirmação interpretados como probabilidades*. Nesse sentido, o teorema de Bayes e o maquinário de probabilidades fornecem uma metodologia a partir da qual relações de confirmação, desconfirmação e neutralidade evidencial são estabelecidas. Se uma proposição evidencial e oferece suporte, embora não conclusivo, a favor de uma certa hipótese h, então e confirma h; se e objeta h, então e desconfirma h; se e não suporta nem objeta h, então e é neutra ou evidencialmente irrelevante sobre h.

Antes de falarmos sobre os conceitos de confirmação e as medidas de diferença, algumas observações precisam ser consideradas. Primeira, embora nem sempre explicitamente declarado, há um princípio que equaliza *graus racionais de crença*, representados por probabilidades subjetivas restringidas pelas condições previamente estabelecidas, com *graus de confirmação*, representados por probabilidades epistêmicas. Conforme o princípio *de evidência a graus de crença*[25], assinalado por D. H. Mellor (2005, p. 80), quanto mais forte e suporta h, maior o grau de crença que um agente deve ter em h condicionado em e. Nessa perspectiva, Bayesianos tipicamente alegam que as relações de suporte evidencial são relativas ao agente e à sua situação doxástica: a concepção de que não há um modo puramente objetivo de suporte evidencial. Segunda, ao analisar a relação de confirmação (ou de desconfirmação ou, ainda, de neutralidade) que e fornece a h, considera-se toda evidência disponível ao agente. Não somente a peça de evidência e, mas todo o *conhecimento de fundo* (*background knowledge*) relativo ao agente é considerado, este último representado pela variável k. Assim, se novas peças de evidência $e_1, e_2, \ldots, e_n$ são obtidas por S, então elas devem ser integradas ao conjunto de estoque de conhecimento k, ou ao conjunto total de informações de fundo, de S. Os graus anteriores de crença, os *priors*, precisam estar condicionados pelo menos em tal estoque de conhecimento ou informações de fundo armazenadas.

Dois conceitos elementares de confirmação podem ser distinguidos: absoluto e incremental. Confirmação absoluta estipula um limiar de valor específico de referência: uma certa evidência suporta fortemente uma determinada hipótese se o limiar de valor χ é superado, desde que $\chi > 0$. Em sentido rigoroso, a definição de confirmação absoluta é formulada nos seguintes termos: e confirma absolutamente h *sse* $Pr(h \mid e \wedge k) > \chi$. Se $\chi = \frac{1}{2}$ e $Pr(h \mid e \wedge k) > \frac{1}{2}$, então $Pr(\neg h \mid e \wedge k) < \frac{1}{2}$. O problema, no entanto, concentra-se na justificação de um determinado valor numérico, e. g. $\frac{1}{2}$, como o limiar de valor apropriado para suporte evidencial absoluto, impedindo que tal limiar de valor seja arbitrariamente fixado. Por isso, Bayesianos têm procurado evitar o conceito de confirmação absoluta em prol do conceito de confirmação incremental.

Confirmação incremental, desconfirmação e neutralidade ou irrelevância evidencial são definidas na devida ordem (Howson e Urbach, 2006, p. 92; Earman, 1992, p. 94):

$$e \text{ confirma } h \textit{ sse } Pr(h \mid e \wedge k) > Pr(h \mid k)$$
$$e \text{ desconfirma } h \textit{ sse } Pr(h \mid e \wedge k) < Pr(h \mid k)$$
$$e \text{ é evidencialmente irrelevante para } h \textit{ sse } Pr(h \mid e \wedge k) = Pr(h \mid k)$$

Se e incrementalmente confirma h, considerando k, então h é menos provável na ausência da evidência e, ou seja, e aumenta a probabilidade de h. De acordo com a avaliação de Joyce (2004, p. 143), o impacto que uma nova evidência e causa sobre uma hipótese h, ao ser adicionada ao estoque de conhecimento k de um agente S, revela uma

[25] Realizamos uma distinção entre probabilidades epistêmicas e *credences* na seção 3. Nem sempre Bayesianos usam tal tipologia, assumindo apenas que confirmação diz respeito a graus subjetivos de crença que satisfazem os axiomas e regras da função probabilística. De todo modo, dado o nosso percurso, parece apropriado explicitar um princípio que proponha uma teoria unificacionista entre graus racionais de crença e graus de confirmação.

discrepância entre o grau de S em h condicionado apenas em k, $Pr(h \mid k)$, e o seu grau em h condicionado em $e \wedge k$, $Pr(h \mid e \wedge k)$, quando e incrementalmente confirma h.

Por último, a teoria Bayesiana de confirmação também é bastante frutífera em relação às medidas de confirmação. Existe, entretanto, uma variedade de medidas de confirmação na literatura sobre Bayesianismo[26]. Não se pretende discuti-las todas aqui. Para o propósito atual, duas delas serão de relevância significante no tocante ao problema da evidência antiga (objeto da próxima seção) *as medidas de diferença d* (Earman, 1992, p. 119) e s (Christensen, 1999, p. 450 e Joyce, 2004, p. 144):

$$d(h,e,k) = Pr(h \mid e \wedge k) - Pr(h \mid k)$$
$$s(h,e,k) = Pr(h \mid e \wedge k) - Pr(h \mid \neg e \wedge k)$$

Se $Pr(h \mid e \wedge k) > Pr(h \mid k)$, então $d(h,e,k) > 0$; se $Pr(h \mid e \wedge k) < Pr(h \mid k)$, então $d(h,e,k) < 0$ e se $Pr(h \mid e \wedge k) = Pr(h \mid k)$, então $d(h,e,k) = 0$. Se $Pr(h \mid e \wedge k) > Pr(h \mid \neg e \wedge k)$, então $s(h,e,k) > 0$; se $Pr(h \mid e \wedge k) < Pr(h \mid \neg e \wedge k)$, então $s(h,e,k) < 0$ e se $Pr(h \mid e \wedge k) = Pr(h \mid \neg e \wedge k)$, então $s(h,e,k) = 0$. Em suma, confirmação se > 0, desconfirmação se < 0 e neutralidade ou irrelevância evidencial se $= 0$.

As *medidas de diferença d* e s captam fatores distintos de impacto da evidência e sobre a hipótese h: a primeira avalia a força que e desempenha sobre h em comparação com a força de h sem e (h condicionado somente em k); a segunda avalia a força que e desempenha sobre h (h condicionado em $e \wedge k$) em comparação com a força que $\neg e$ exerce sobre h (h condicionado em $\neg e \wedge k$). Como veremos, embora a primeira seja uma medida bastante canônica, ela parece ser problemática e ineficiente na resolução do problema da evidência antiga.

Problemas e Objeções

Alguns dos principais problemas e objeções serão tratados nesta seção. Em especial, o problema dos *priors* e o problema da evidência antiga impõem sérios desafios à proposta do probabilismo e Bayesianismo em epistemologia.

Primeiro, uma objeção diz que agentes ordinários nem sempre são capazes de satisfazer o modelo de racionalidade proposto pelo probabilismo e Bayesianismo. Na verdade, há uma suposição conhecida como *suposição da onisciência lógica* (Garber, 1983) que, advertidamente, parece exigir um padrão muito alto, inatingível, a agentes ordinários. Agentes precisariam satisfazer as regras e o axiomas do maquinário de probabilidades, sendo coerentes probabilisticamente, e todas consequências lógicas que se

[26] Várias delas disponíveis em Huber (2010, sec. 6) e Fitelson (1999, sec. 1). Por exemplo, a medida de razão dos *likelihoods*, $\frac{Pr(e|h \wedge k)}{Pr(e|\neg h \wedge k)}$, e a medida dos *likelihoods* normalizada, razão logarítimica dos *likelihoods*, $\log[\frac{Pr(e|h \wedge k)}{Pr(e|\neg h \wedge k)}]$.

seguem de tal maquinário: *e.g.* deve-se atribuir valor maximal 1 a tautologias e valor minimal 0 a inconsistências e contradições lógicas; se p acarreta q, a probabilidade de q não pode ser menor do que a de p; entre outras. A alegação é de que são exigências muito altas para que agentes as satisfaçam em todas circunstâncias ordinárias de suas práticas epistêmicas. Presumivelmente, uma resposta possível endereçada a tal objeção sustenta que os padrões exigidos pelo modelo formal do probabilismo e Bayesianismo devem ser tomados como *ideais*. São, rigorosamente, padrões para *agentes idealmente racionais*.

Segundo, duas objeções podem ser levantadas contra o argumento do *Dutch Book*. Para Pollock e Cruz (1999, p. 95), o *Dutch Book* não diz respeito à racionalidade epistêmica. Em última instância, o que estaria em jogo no caso de um agente não violar o aparato formal de probabilidade concerne à racionalidade prática, a saber, o agente é racional do ponto de vista *prático* se não aceita apostas que o conduzem à perda garantida. Ele não empreende uma ação que o coloca numa situação praticamente indesejável. Isso, no entanto, não o faz necessariamente racional do ponto de vista *epistêmico*[27]. Nessa perspectiva, o *Dutch Book* no máximo revelaria incoerência probabilística nos quocientes de aposta do agente. Uma outra crítica decisiva ao *Dutch Book* é surpreendentemente formulada por Hájek (2009, p. 232). Há um argumento conhecido como *Czech Book* que oferece boas razões contra o probabilismo. Se os graus de crença, representados por quocientes de apostas, de um agente *S* não satisfazem os axiomas e regras de probabilidade, então existe um conjunto de apostas que assegura ganho a *S*. Se, por outro lado, os graus de crença de *S* satisfazem os axiomas e regras de probabilidade, então não existe um conjunto de apostas que assegura ganho a *S*. Assim, parece ser um *desideratum* para o agente não satisfazer o maquinário de probabilidades: ser coerente probabilisticamente não oferece ganho a *S*. Tal argumento é genuinamente um *plot twist*, Hájek (2009, p. 231) o denomina de *diabolical twist*, nas pretensões de um defensor do probabilismo.

Terceiro, o problema dos *priors*. Pode haver desacordo se graus anteriores de crença, representados formalmente pelos *priors*, são ou não determinados por condições restritivas adicionais. Vamos supor que não temos nenhuma evidência relevante sobre um determinado mecanismo de jogadas de moeda, somos ignorante se ele é justo ou viciado num determinado resultado, com exceção que sabemos dos resultados possíveis de ou cara ou coroa, o que configura o *espaço amostral ou de possibilidades* constituído num conjunto tal que $\Omega = \{cara, coroa\}$. Devemos distribuir de maneira equânime os valores de probabilidade entre os dois resultados possíveis? Por um lado, o Bayesianismo objetivo[28] pode invocar o princípio de indiferença[29] de tal modo que devemos, além de satisfazer as

[27] Uma tentativa de *despragmatização* do argumento do *Dutch Book* é encontrada em Christensen (2004, p. 116-124).

[28] Jon Williamson defende uma forma robusta de Bayesianismo objetivo em *In Defence of Objective Bayesianism* (2010).

[29] O princípio de indiferença diz que se um agente *S* tem unicamente a evidência de que um *espaço de amos-*

outras restrições e regras da função probabilística, atribuir o valor de $\frac{1}{n}$ a cada resultado de um total de n possibilidades. Assim, o princípio de indiferença funcionaria como um critério *a priori* de atribuição de valores em relação aos *priors* e, alegadamente, seríamos racionais ao aplicar $\frac{1}{2}$ a cada um dos resultados do nosso exemplo. Por outro lado, o Bayesianismo subjetivo[30], talvez numa forma mais radical, alega que simplesmente não temos restrições adicionais sobre os graus anteriores de crença. Devemos ser coerentes probabilisticamente, mas não há nada que nos autorize a distribuir graus equânimes entre cada um dos resultados. Assim, não há uma atribuição correta de um valor específico entre $[0,1]$ em tal cenário. A despeito da própria dificuldade de justificar algum princípio sobre os *priors*, como é o caso do princípio de indiferença, a questão exposta aqui pode separar Bayesianos subjetivos dos Bayesianos objetivos. Mas, como lembra W. Talbott (2008), em um cenário com nenhuma evidência sequer disponível ao agente, nem mesmo algum dado sobre o *espaço de possibilidades*, pode ser o caso que Bayesianos concordem sobre os *priors*. Apesar de acordo *minimum*, se não há evidência ou informação disponível, então não há base nenhuma sobre a qual graus anteriores de crença de um agente podem se apoiar.

Quarto, o problema da evidência antiga, talvez o desafio mais contundente ao Bayesianismo (Glymour, 1980, p. 85-93; Earman, 1992, p. 119-135). Suponha num dado tempo t que uma evidência antiga e suporta uma hipótese h. Suponha que em t' descobre-se uma relação lógica entre ambas, a saber, h acarreta e. Assumindo, além disso, que $Pr(e) = 1$ em t'. Se h acarreta e, que é o caso em t', então $h \equiv (h \wedge e)$. Se $h \equiv (h \wedge e)$, então $Pr(h) = Pr(h \wedge e)$ (**Teorema §1** disponível no Apêndice). Assim, pela definição de probabilidade condicional, $Pr(e \mid h) = \frac{Pr(h \wedge e)}{Pr(h)}$. Por substituição, $Pr(e \mid h) = \frac{Pr(h)}{Pr(h)}$. Portanto, $Pr(e \mid h) = 1$, considerando que $Pr(h) \neq 0$. Usando uma variação de $\langle Bayes\ \S1 \rangle$, $Pr(h \mid e) = \frac{Pr(e \mid h) \times Pr(h)}{Pr(e)}$ (disponível no **Teorema §5**), dado que $Pr(e \mid h) = 1$ e $Pr(e) = 1$ em t', conclui-se que $Pr(h \mid e) = Pr(h)$ em t'. Se usamos a medida de diferença $d(h,e) = Pr(h \mid e) - Pr(h)$, suprimindo a variável k no condicional, temos $d(h,e) = 0$. Portanto, a evidência antiga e não desempenha nenhum papel de suporte de confirmação incremental sobre h em t'. Em outros termos, a medida de diferença d mostra que e é evidencialmente irrelevante para h em t'. Suprimida a variável k, a medida de confirmação $s(h,e) = Pr(h \mid e) - Pr(h \mid \neg e)$, proposta por Christensen[31] (1999) e Joyce

tra Ω tem n possibilidades mutuamente exclusivas e conjuntamente exaustivas, $\Omega = \{h_1, h_2, \cdots, h_n\}$, então é racional para S atribuir grau de $\frac{1}{n}$ a cada uma de tais possibilidades de Ω. Mais informações sobre o princípio de indiferença em Howson (2009, sec. 1).

[30] Uma versão de Bayesianismo subjetivo é oferecida por Richard Jeffrey em *Subjective Probability: The Real Thing* (2004).

[31] Christensen propõe uma versão alternativa do problema da evidência antiga em que $Pr(e) \approx 1$. A questão, no entanto, é que o problema parece se perder se assumimos que $Pr(e) \neq 1$, como demonstra Earman (1992, p. 121). Em última instância, Christensen (1999, p. 459) mostra-se bastante cético sobre s

(2004), pode ser obtida da medida d. Divide-se a medida d, como Christensen (1999, p. 450) demonstra em uma nota de rodapé, por $Pr(\neg e)$. Desse modo, $\frac{Pr(h|e)-Pr(h)}{Pr(\neg e)}$ (l). Dado que $h \equiv [(h \wedge e) \vee (h \wedge \neg e)]$, $Pr(h) = Pr[(h \wedge e) \vee (h \wedge \neg e)]$ (pelo **Teorema §1**). $Pr[(h \wedge e) \vee (h \wedge \neg e)] = Pr(h \wedge e) + Pr(h \wedge \neg e)$ porque $h \wedge \neg e$ e $h \wedge e$ são mutuamente exclusivas (pelo axioma de aditividade finita). Por substituição, $Pr(h) = Pr(h \wedge e) + Pr(h \wedge \neg e)$. Por seu turno, $Pr(h \wedge e) = Pr(h \mid e) \times Pr(e)$ e $Pr(h \wedge \neg e) = Pr(h \mid \neg e) \times Pr(\neg e)$ (pela definição de probabilidade condicional). Assim, $Pr(h) = Pr(h \mid e) \times Pr(e) + Pr(h \mid \neg e) \times Pr(\neg e)$. Substituindo em (l), $\frac{Pr(h|e) \ - \ [Pr(h|e) \times Pr(e) \ + \ Pr(h|\neg e) \times Pr(\neg e)]}{Pr(\neg e)}$. Assim, fazendo uma alteração no numerador, $\frac{[Pr(h|e) \ - \ Pr(h|e) \times Pr(e)] \ - \ Pr(h|\neg e) \times Pr(\neg e)}{Pr(\neg e)}$. Por uma regra de fatoração, $(a - a \times b) = a \times (1 - b)$, consegue-se $\frac{Pr(h|e) \times [1 - Pr(e)] \ - \ Pr(h|\neg e) \times Pr(\neg e)}{Pr(\neg e)}$. Uma vez que $Pr(\neg e) = 1 - Pr(e)$, $\frac{Pr(h|e) \times Pr(\neg e) \ - \ Pr(h|\neg e) \times Pr(\neg e)}{Pr(\neg e)}$. Portanto, $Pr(h \mid e) - Pr(h \mid \neg e)$, que é a medida de confirmação s. Tal medida seria uma alternativa na resolução do problema da evidência antiga, uma vez que o resultado da medida s ainda pode ser positivo, ao passo que d tem valor minimal 0. Não obstante, parece correto afirmar, tal como Earman (1992, p. 121) o faz, que s não é uma medida de confirmação apropriada. Se mantivermos as condições anteriores de que h acarreta e, $Pr(h \mid e) = Pr(h)$ e $Pr(e) = 1$ em t', temos $s(h,e) = Pr(h) - Pr(h \mid \neg e)$. A situação agrava-se ainda mais no que diz respeito à medida de confirmação s porque, ao fim e ao cabo, se $Pr(h \mid \neg e) = 0$ em t', então $s(h,e) = Pr(h)$. Nesse sentido, a evidência e não desempenharia nenhum papel de confirmação de h, ainda que o valor da medida s possa ser positivo, ou seja, mesmo que $s(h,e) > 1$. Parece ser o caso que s é, de fato, uma medida inapropriada. Considerando as condições estabelecidas anteriormente, temos $Pr(e) = 1$ e $Pr(e \mid h) = 1$ em t'. Se $Pr(\neg e) = 1 - Pr(e)$ e $Pr(e) = 1$, então $Pr(\neg e) = 0$. Se $Pr(e \mid h) = 1$, então $Pr(\neg e \mid h) = 0$ porque $Pr(\neg e \mid h) = 1 - Pr(e \mid h)$. Por sua vez, $Pr(h \mid \neg e) = \frac{Pr(h \ \wedge \ \neg e)}{Pr(\neg e)}$. Se, além de $Pr(\neg e) = 0$, $Pr(h \wedge \neg e) = 0$, então $Pr(h \mid \neg e) = \frac{0}{0}$. Há boas razões, entretanto, para considerar que o valor de uma divisão de *zero* por *zero* é matematicamente indeterminada. Suponha que $x \times 0 = 0$ tal que $\forall x \in \Re$. Se $\frac{0}{0} = a$ e, multiplicando 0 nos dois lados da equação, $\frac{0}{0} \times 0 = a \times 0$, consegue-se $0 = a \times 0$, sendo que a pode ser qualquer número que pertence ao conjunto dos números reais. Portanto, a é indeterminado.

Considerações Finais

Probabilismo e Bayesianismo propõem uma teoria formalmente elegante, combinando aspectos *sincrônicos* e *diacrônicos* sobre o modelo de graus de crença para agentes doxásticos. Considerações e distinções conceituais foram inicialmente expostas, pro-

ser uma medida de confirmação probabilística apropriada. A propósito, Eells (1990, p. 207) apresenta uma sistematização bastante útil sobre versões diferentes do problema da evidência antiga. Aqui apenas oferecemos o que parece ser a versão *default* do problema.

curando fornecer os elementos básicos do *calculus* probabilístico e caracterizando a interpretação subjetiva da função de probabilidades. Após a explicação da relação entre quocientes de apostas e graus de crença, o núcleo da teoria foi propriamente alcançado: o *Dutch Book*, o princípio principal, os princípios de *condicionalização* §1 e §2, o teorema de Bayes e a teoria de confirmação Bayesiana. Vimos, todavia, que o probabilismo e o Bayesianismo não estão imunes a severas críticas. O problema da evidência antiga é verdadeiramente o *calcanhar de Aquiles* do Bayesianismo, sobretudo em relação às medidas de confirmação, embora nem todas as medidas tenham sido avaliadas neste artigo. Em todo caso, respostas mais convincentes aos problemas formulados precisam ser empreendidas pelo probabilismo e o Bayesianismo se estes pretendem ser uma teoria robusta em epistemologia.

Apêndice: Demonstração de Teoremas

Equivalências lógicas:
(a) $p \equiv [(p \wedge q) \vee (p \wedge \neg q)]$
(b) $q \equiv [(p \wedge q) \vee (\neg p \wedge q)]$
(c) $(p \vee q) \equiv [(p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge q)]$

Teorema §1: Se $p \equiv q$, então $Pr(p) = Pr(q)$.
Prova: se $p \equiv q$, então $p \vee \neg q$ é uma tautologia. Assim, pelo axioma (2'), $Pr(p \vee \neg q) = 1$. Pelo axioma (3) de aditividade finita, $Pr(p) + Pr(\neg q) = Pr(p \vee \neg q)$, pois p e $\neg q$ são mutuamente exclusivas, uma vez que $p \equiv q$. Por conseguinte, $Pr(p) + Pr(\neg q) = 1$. Pelo axioma (2'), $Pr(q \vee \neg q) = 1$, pois $q \vee \neg q$ é uma tautologia. Assim, pelo axioma (3), $Pr(q \vee \neg q) = Pr(q) + Pr(\neg q)$, pois q e $\neg q$ são mutuamente exclusivas. Por conseguinte, $Pr(q) + Pr(\neg q) = 1$. Deriva-se da fórmula anterior que $Pr(\neg q) = 1 - Pr(q)$. Substituindo em $Pr(p) + Pr(\neg q) = 1$, segue-se que $Pr(p) + 1 - Pr(q) = 1$. Portanto, $Pr(p) = Pr(q)$.

Teorema §2: $Pr(p \vee q) = Pr(p) + Pr(q) - Pr(p \wedge q)$.
Prova: pelo teorema §1 e equivalência (a), $Pr(p) = Pr[(p \wedge q) \vee (p \wedge \neg q)]$. Pelo axioma (3), $Pr[(p \wedge q) \vee (p \wedge \neg q)] = Pr(p \wedge q) + Pr(p \wedge \neg q)$, pois $p \wedge q$ e $p \wedge \neg q$ são mutuamente exclusivas. Assim, $Pr(p) = Pr(p \wedge q) + Pr(p \wedge \neg q)$ (z). Pelo teorema §1 e equivalência (b), $Pr(q) = Pr[(p \wedge q) \vee (\neg p \wedge q)]$. Uma vez que $p \wedge q$ e $\neg p \wedge q$ são mutuamente exclusivas, segue-se que $Pr[(p \wedge q) \vee (\neg p \wedge q)] = Pr(p \wedge q) + Pr(\neg p \wedge q)$ pelo axioma (3). Por conseguinte, $Pr(q) = Pr(p \wedge q) + Pr(\neg p \wedge q)$ (s). Pelo teorema §1 e equivalência (c), $Pr(p \vee q) = Pr[(p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge q)]$. Uma vez que $p \wedge q$, $p \wedge \neg q$ e $\neg p \wedge q$ são mutuamente exclusivas, segue-se que $Pr[(p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge q)] = Pr(p \wedge q) + Pr(p \wedge \neg q) + Pr(\neg p \wedge q)$ pelo axioma (3). Assim, $Pr(p \vee q) = Pr(p \wedge q) + Pr(p \wedge \neg q) + Pr(\neg p \wedge q)$ (r). Consegue-se de (z) que $Pr(p \wedge q) = Pr(p) - Pr(p \wedge \neg q)$. Consegue-se de (s) que $Pr(\neg p \wedge q) = Pr(q) - Pr(p \wedge q)$. Substituindo em (r), segue-se que $Pr(p \vee q) = Pr(p) - Pr(p \wedge \neg q) + Pr(q) - Pr(p \wedge q) + Pr(p \wedge \neg q)$. Portanto, $Pr(p \vee q) = Pr(p) + Pr(q) - Pr(p \wedge q)$.

Teorema §3: Se $p \vDash q$, então $Pr(p) \leq Pr(q)$.
Prova: se $p \vDash q$, então $p \equiv (p \wedge q)$. Assumindo que $p \vDash q$, $Pr(p) = Pr(p \wedge q)$ (n) pelo teorema §1 . Assim, pela equivalência (b) e pelo teorema §1, $Pr(q) = Pr[(p \wedge q) \vee (\neg p \wedge q)]$. Pelo axioma (3), $Pr[(p \wedge q) \vee (\neg p \wedge q)] = Pr(p \wedge q) + Pr(\neg p \wedge q)$, pois $p \wedge q$ e $\neg p \wedge q$ são mutuamente exclusivas. Segue-se, assim, que $Pr(q) = Pr(p \wedge q) + Pr(\neg p \wedge q)$. Por conseguinte, usando (n), segue-se que $Pr(q) = Pr(p) + Pr(\neg p \wedge q)$. Se $Pr(\neg p \wedge q) = 0$, então $Pr(q) = Pr(p)$. Se $Pr(\neg p \wedge q) > 0$, então $Pr(q) > Pr(p)$. Portanto, $Pr(p) \leq Pr(q)$.

Teorema §4: $Pr(p) = [Pr(q) \times Pr(p \mid q)] + [Pr(\neg q) \times Pr(p \mid \neg q)]$, dado que $1 > Pr(q) > 0$.

Prova: Suposição de que $1 > Pr(q) > 0$. Pelo teorema §1 e equivalência (a), segue-se que $Pr(p) = Pr[(p \wedge q) \vee (p \wedge \neg q)]$. Pelo axioma (3), $Pr[(p \wedge q) \vee (p \wedge \neg q)] = Pr(p \wedge q) + Pr(p \wedge \neg q)$, pois $p \wedge q$ e $p \wedge \neg q$ são mutuamente exclusivas. Assim, $Pr(p) = Pr(p \wedge q) + Pr(p \wedge \neg q)$ (g). Pela definição de probabilidade condicional, $Pr(p \mid q) = \frac{Pr(p \wedge q)}{Pr(q)}$. Por conseguinte, $Pr(p \wedge q) = Pr(p \mid q) \times Pr(q)$. $Pr(p \mid \neg q) = \frac{Pr(p \wedge \neg q)}{Pr(\neg q)}$ pela definição de probabilidade condicional. Assim, $Pr(p \wedge \neg q) = Pr(p \mid \neg q) \times Pr(\neg q)$. Portanto, substituindo em (g), $Pr(p) = [Pr(q) \times Pr(p \mid q)] + [Pr(\neg q) \times Pr(p \mid \neg q)]$.

Teorema §5:

$$Pr(h \mid e) = \frac{Pr(h) \times Pr(e \mid h)}{[Pr(h) \times Pr(e \mid h)] + [Pr(\neg h) \times Pr(e \mid \neg h)]}$$

dado que $Pr(e) > 0$.

Prova: Suposição de que $Pr(e) > 0$. $Pr(h \mid e) = \frac{Pr(h \wedge e)}{Pr(e)}$ e $Pr(e \mid h) = \frac{Pr(e \wedge h)}{Pr(h)}$ pela definição de probabilidade condicional. Assim, $Pr(e \wedge h) = Pr(e \mid h) \times Pr(h)$. Por conseguinte, uma vez que $h \wedge e$ e $e \wedge h$ são logicamente equivalentes, $Pr(h \mid e) = \frac{Pr(e \mid h) \times Pr(h)}{Pr(e)}$. Pelo teorema §4, $Pr(e) = [Pr(h) \times Pr(e \mid h)] + [Pr(\neg h) \times Pr(e \mid \neg h)]$. Portanto, depreende-se que $Pr(h \mid e) = \frac{Pr(h) \times Pr(e \mid h)}{[Pr(h) \times Pr(e \mid h)] + [Pr(\neg h) \times Pr(e \mid \neg h)]}$.

Teorema §6:

$$Pr(h \mid e \wedge k) = \frac{Pr(h \mid k) \times Pr(e \mid h \wedge k)}{Pr(e \mid k)}$$

dado que $Pr(e \mid k) > 0$.

Prova: Suposição de que $Pr(e \mid k) > 0$. $Pr(h \mid e \wedge k) = \frac{Pr(h \wedge e \wedge k)}{Pr(e \wedge k)}$ e $Pr(e \mid h \wedge k) = \frac{Pr(e \wedge h \wedge k)}{Pr(h \wedge k)}$ pela definição de probabilidade condicional. Por conseguinte, $Pr(e \wedge h \wedge k) = Pr(e \mid h \wedge k) \times Pr(h \wedge k)$. Assim, uma vez que $e \wedge h \wedge k$ e $h \wedge e \wedge k$ são equivalentes, $Pr(h \mid e \wedge k) = \frac{Pr(e \mid h \wedge k) \times Pr(h \wedge k)}{Pr(e \wedge k)}$. $Pr(h \mid k) = \frac{Pr(h \wedge k)}{Pr(k)}$ e $Pr(e \mid k) = \frac{Pr(e \wedge k)}{Pr(k)}$ pela definição de probabilidade condicional. Assim, $Pr(h \wedge k) = Pr(h \mid k) \times Pr(k)$ e $Pr(e \wedge k) = Pr(e \mid k) \times Pr(k)$. Por conseguinte, $Pr(h \mid e \wedge k) = \frac{Pr(e \mid h \wedge k) \times Pr(h \mid k) \times Pr(k)}{Pr(e \mid k) \times Pr(k)}$. Portanto, $Pr(h \mid e \wedge k) = \frac{Pr(h \mid k) \times Pr(e \mid h \wedge k)}{Pr(e \mid k)}$.

Referências

BAYES, T. **An Essay towards solving a Problem in the Doctrine of Chances**. In: SWINBURNE, R. (Ed.). **Bayes's Theorem**. Oxford: Oxford University Press. Proceedings of the British Academy 113, 2002. p. 117–149.

BOVENS, L.; HARTMANN, S. **Bayesian Epistemology**. Oxford: Oxford University Press, 2003.

BRADLEY, D. **A Critical Introduction to Formal Epistemology**. London: Bloomsbury, 2015.

CARNAP, R. **Logical Foundations of Probability**. 2nd ed. Chicago: The University of Chicago Press, 1962.

CHRISTENSEN, D. Measuring confirmation. **The Journal of Philosophy**, v. 96, n. 09, p. 437–461, 1999.

CHRISTENSEN, D. **Putting Logic in its Place: Formal Constraints on Rational Belief**. New York: Oxford University Press, 2004.

EARMAN, J. **Bayes or Bust?** Cambridge, MA: MIT Press, 1992.

EELLS, E. Bayesian problems of old evidence. **Scientific Theories, Minnesota Studies in the Philosophy of Science**, v. 10, p. 205–223, 1990.

FINETTI, B. D. **Foresight: Its Logical Laws, Its Subjective Sources**. In: KYBURG, H.; SMOKLER, H. (Ed.). **Studies in Subjective Probability**. New York: Wiley, 1964. p. 93–159.

FITELSON, B. The plurality of bayesian measures of confirmation and the problem of measure sensitivity. **Philosophy of Science**, v. 66, Suplement, p. S362–S378, 1999.

GARBER, D. Old evidence and logical omniscience in bayesian confirmation theory. **Testing Scientific Theories. Midwest Studies in the Philosophy of Science**, University of Minnesota Press, v. 10, p. 99–131, 1983.

GILLIES, D. **Philosophical Theories of Probability**. London: Routledge, 2000.

GLYMOUR, C. **Theory and Evidence**. Princeton: Princeton University Press, 1980.

HACKING, I. **An Introduction to Probability and Inductive Logic**. Cambridge: Cambridge University Press, 2001.

HÁJEK, A. **Arguments for – or against – Probabilism?** In: HUBER, F.; SCHMIDT-PETRI, C. (Ed.). **Synthese Library 342, Degrees of Belief**. Dordrecht: Synthese, 2009. p. 229–251.

HÁJEK, A. **Interpretations of Probability**. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. http://plato.stanford.edu/entries/probability-interpret/: (Fall/Winter), 2011.

HÁJEK, A.; HARTMANN, S. **Bayesian Epistemology**. In: DANCY, J.; SOSA, E.; STEUP, M. (Ed.). **A Companion to Epistemology**. 2nd ed. Malden: Wiley-Blackwell, 2010. p. 93–105.

HARTMANN, S.; SPRENGER, J. **Bayesian Epistemology**. In: BERNECKER, S.; PRITCHARD, D. (Ed.). **The Routledge Companion to Epistemology**. London and New York: Routledge, 2011. p. 609–620.

HOWSON, C. **Epistemic Probability and Coherent Degrees of Belief**. In: HUBER, F.; SCHMIDT-PETRI, C. (Ed.). **Synthese Library 342, Degrees of Belief**. Dordrecht: Synthese, 2009. p. 97–119.

HOWSON, C.; URBACH, P. **Scientific Reasoning: the Bayesian Approach**. 3rd ed. Chicago and La Salle: Open Court Publishing, 2006.

HUBER, F. **Belief and Degrees of Belief**. In: HUBER, F.; SCHMIDT-PETRI, C. (Ed.). **Synthese Library 342, Degrees of Belief**. Dordrecht: Synthese, 2009. p. 1–33.

HUBER, F. **Confirmation and Induction**. In: FIESER, J.; DOWDEN, B. (Ed.). **Internet Encyclopedia of Philosophy**. http://www.iep.utm.edu/conf-ind/: (disponível online), 2010.

JEFFREY, R. C. **The Logic of Decision**. Chicago and London: The University of Chicago Press, 1983.

JEFFREY, R. C. **Subjective Probability: The Real Thing**. Cambridge: Cambridge University Press, 2004.

JOYCE, J. A nonpragmatic vindication of probabilism. **Philosophy of Science**, v. 65, p. 575–603, 1998.

JOYCE, J. **Bayesianism**. In: MELE, A.; RAWLING, P. (Ed.). **The Oxford Handbook of Rationality**. New York: Oxford University Press, 2004. p. 132–155.

KEMENY, J. G. Fair bets and inductive probabilities. **The Journal of Symbolic Logic**, v. 20, n. 03, p. 263–273, 1955.

KOLMOGOROV, A. N. **Foundations of the Theory of Probability**. 2nd ed. New York: Chelsea Publishing Company, 1956.

LEWIS, D. **A Subjectivist's Guide to Objective Chance**. In: JEFFREY, R. C. (Ed.). **Studies in Inductive Logic, vol. II**. Berkeley: University of California Press, 1980. p. 263–293.

MAHER, P. **Betting on Theories**. Cambridge: Cambridge University Press, 1993.

MELLOR, D. H. **Probability: A Philosophical Introduction**. London: Routledge, 2005.

POLLOCK, J. L.; CRUZ, J. **Contemporary Theories of Knowledge**. 2nd ed. Lanham, MD: Rowman & Littlefield, 1999.

RAMSEY, F. P. **Truth and Probability**. In: RAMSEY, F. P.; BRAITHWAITE, R. B. (Ed.). **The Foundations of Mathematics and Other Logical Essays**. London: Routledge, 1950. p. 156–198.

TALBOTT, W. **Bayesian Epistemology**. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. http://plato.stanford.edu/entries/epistemology-bayesian/: (Spring/Summer), 2008.

VINEBERG, S. **Dutch Book Arguments**. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. http://plato.stanford.edu/entries/dutch-book/: (Spring/Summer), 2011.

WEISBERG, J. **Formal Epistemology**. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. http://plato.stanford.edu/entries/formal-epistemology/: (Spring/Summer), 2015.

WILLIAMSON, J. **In Defence of Objective Bayesianism**. Oxford: Oxford University Press, 2010.