PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL

FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS

PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

ANDRÉ LUIZ DE ALMEIDA LISBÔA NEIVA

PROBABILIDADE E BAYESIANISMO NA TEORIA EPISTÊMICA DE RICHARD SWINBURNE

PORTO ALEGRE

2016

ANDRÉ LUIZ DE ALMEIDA LISBÔA NEIVA

PROBABILIDADE E BAYESIANISMO NA TEORIA EPISTÊMICA DE RICHARD SWINBURNE

Dissertação apresentada como requisito para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Filosofia da Faculdade de Filosofia e Ciências Humanas da Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dr. Roberto Hofmeister Pich

PORTO ALEGRE

2016

ANDRÉ LUIZ DE ALMEIDA LISBÔA NEIVA

PROBABILIDADE E BAYESIANISMO NA TEORIA EPISTÊMICA DE RICHARD SWINBURNE

Dissertação apresentada como requisito para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Filosofia da Faculdade de Filosofia e Ciências Humanas da Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dr. Roberto Hofmeister Pich

Aprovada em:______de___________de 2016.

BANCA EXAMINADORA:

Prof. Dr. Roberto Hofmeister Pich (Orientador) – PUCRS

Prof. Dr. Luis Fernando Munaretti da Rosa – LMU Munich (MCMP)

Prof. Dr. Rodrigo Borges – UNICAMP (FAPESP)

PORTO ALEGRE

2016

AGRADECIMENTOS

Gostaria de agradecer ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por ter custeado a minha pesquisa com uma bolsa integral e ao Programa de Pós-Graduação em Filosofia da PUCRS pela ótima estrutura que tive a oportunidade de usufruir.

Agradecimento especial ao meu orientador, Prof. Dr. Roberto Hofmeister Pich, que me acompanhou nos dois anos de mestrado e com quem muito aprendi durante esse período. Muito obrigado pela confiança, por ter me proporcionado organizar e participar de um evento tão importante como o 'Colóquio Richard Swinburne' no ano de 2015, no qual tive a oportunidade de discutir argumentos com o próprio Professor Richard Swinburne, e pela excelente e cuidadosa orientação.

Aos meus amigos e colegas do Grupo de Epistemologia Formal, que iniciou no ano de 2014, João Fett, Luiz Paulo Cichoski, Leonardo Ruivo, Thiago Santin, Lucas Roisenberg, Felipe Miguel, Rossul Padilha, Ricardo Rangel, Valentinne Serpa, Gregory Gaboardi e, em especial, ao Luis Rosa, grande parceiro com quem tive o prazer de conviver, aprender muito e quem originalmente teve a ideia do nosso Grupo de Pesquisa. Além disso, agradeço ao Danilo Dantas pelos comentários e sugestões na versão preliminar da dissertação. Obrigado a todos pelo debate, sugestões, argumentos e refutações nos bastidores da Universidade.

A todos os Professores, amigos e colegas que convivi nesses dois anos, particularmente aos professores Cláudio Almeida, Felipe Müller, Nythamar de Oliveira, Agemir Bavaresco, Norman Madarasz, Rogel Oliveira, Kátia Etcheverry, Fabricio Pontin e aos meus amigos Jair Tauchen, Vanessa Nicola Labrea, Jerônimo Milone, Marco Antonio Scapini, Renata Floriano, Evandro Pontel, Bruna Bortolini, Manuela de Mattos e tantos outros.

À minha família e aos meus pais, José Luiz Lisbôa Neiva e Ana Beatriz Moura de Almeida, que sempre me incentivaram do início ao fim.

À minha namorada e companheira Tatiane Marks, pela companhia, apoio incondicional nos estudos, pela paciência quando eu me empolgava com teoremas e parafernálias formais. Muito obrigado!

What's a Bayesian?
Well, I'm one, for example.
(Richard Jeffrey, 1983a)

I confess that I am a Bayesian—
at least I am on Mondays, Wednesdays,
and Fridays.
(John Earman, 1992)

RESUMO

Este trabalho pretende explorar os aspectos centrais da teoria de probabilidade e do Bayesianismo objetivo de Richard Swinburne. No capítulo inaugural, apresentamos e avaliamos alguns pressupostos básicos da teoria de Swinburne, sobretudo as relações entre os conceitos de crença e probabilidade e a sua defesa do contrastivismo doxástico. No capítulo seguinte, oferecemos os axiomas e as definições do maquinário formal do cálculo de probabilidades e, em seguida, examinamos mais minuciosamente os diferentes tipos de probabilidade e o seu emprego no contexto de disputa entre internalismo e externalismo. No último capítulo, analisamos o Bayesianismo de Swinburne e os seus critérios de probabilidade lógica. O critério *a priori* de simplicidade é o mais importante na sua versão de Bayesianismo. Na parte final, discutimos alguns problemas e objeções à sua teoria e ao Bayesianismo como um todo. O problema da evidência antiga continua a ser o principal desafio à teoria de confirmação Bayesiana. Além disso, concluímos que o critério de simplicidade não é bem-sucedido em relação ao *desideratum* pretendido por Swinburne.

Palavras-chave: Bayesianismo. Confirmação. Evidência Antiga. Probabilidade. Simplicidade.

ABSTRACT

This work aims to explore the main aspects of Richard Swinburne's theory of probability and objective Bayesianism. In the first chapter, we present and evaluate some basic assumptions of Swinburne's theory, especially the relations between the concepts of belief and probability and his defense of doxastic contrastivism. In the intermediate chapter, we provide the axioms and definitions of the formal machinery of probability and then we examine more carefully different kinds of probability and its use in the context of dispute between internalism and externalism. In the last chapter, we analyze Swinburne's Bayesianism and his criteria of logical probability. The *a priori* criterion of simplicity is the most important in his version of Bayesianism. In the final part, we discuss some problems and objections to his theory and to Bayesianism in general. The problem of old evidence remains the main challenge to the Bayesian confirmation theory. Furthermore, we conclude that the criterion of simplicity is unsuccessful in achieving Swinburne's purposes.

Keywords: Bayesianism. Confirmation. Old Evidence. Probability. Simplicity.

Sumário

1 Introdução

Probabilidade e Bayesianismo têm ganhado importância em diversas áreas das ciências e da filosofia: estatística, teoria da decisão, teoria dos jogos, filosofia da ciência, epistemologia, filosofia da religião, entre outras. Particularmente em epistemologia, o interesse pelo uso de técnicas e métodos formais do cálculo de probabilidades tem aumentado consideravelmente nos últimos anos. Conceitos e problemas clássicos de epistemologia podem ser analisados mais acuradamente com o uso de ferramentas que as teorias de probabilidade *lato sensu* oferecem com suas metodologias: o cálculo fornece um modelo formal para graus de crença, graus de coerência podem ser explicados por meio de medidas de probabilidade, uma abordagem formal sobre os conceitos de confiabilidade e testemunho permitem uma análise mais refinada dos seus aspectos, pode-se fazer uso de um conceito falibilista de justificação epistêmica cujas bases encontram-se na noção de suporte evidencial probabilístico, o Bayesianismo contribui sobre problemas envolvendo confirmação e indução, e assim por diante.

Entretanto, existem tipos e interpretações distintas sobre o conceito de probabilidade. Os diversos usos dos termos 'probabilidade' e 'provável' não espelham um único tipo geral. Em certos casos, alegações que empregam esses termos designam alguma propriedade física do mundo, como uma propensão natural de um evento ocorrer, dado um conjunto de estados de coisas e certas condições satisfeitas. Por exemplo, 'é provável que um átomo de rádio tenha decaimento radioativo em um certo intervalo de tempo' corresponde a esse tipo de emprego. Por vezes, no entanto, a descrição do que é provável simplesmente refere-se a uma *ratio*, ou a uma proporção, na qual um certo tipo de evento ocorre em uma longa sequência de repetições; por exemplo, 'é tão provável que um mecanismo de jogadas de moeda produza resultado de *coroa* quanto resultado de *cara* em uma sequência de repetições'. Nesse sentido, a *ratio* de uma moeda sair *coroa* (ou *cara*) tende a ser $\frac{1}{2}$. Mas em muitas situações estamos nos referindo ao grau de suporte que um conjunto de evidências total fornece a uma hipótese, isto é, o quão provável ela é, supondo que certas evidências são o caso. Por exemplo, 'supondo o conjunto de evidências astronômicas disponíveis, é provável que a hipótese do *Big Bang* seja verdadeira' concerne a esse terceiro sentido do termo 'provável'.

Ademais, o Bayesianismo não está limitado única e exclusivamente ao teorema de Bayes. Certamente ele exerce papel central nas metodologias Bayesianas. Porém, o termo 'Bayesianismo' refere-se a um conjunto mais abrangente de métodos, princípios e restrições do que somente a aplicação de tal ferramenta matemática. Em primeiro lugar, existem diferentes versões de Bayesianismo. Com uma distinção muito ampla, mas útil, podemos separar Bayesianos em dois grandes grupos: Bayesianos subjetivos e objetivos. Por um lado, Bayesianos subjetivos defendem que graus subjetivos de crença, ou probabilidades subjetivas, devem satisfazer o cálculo probabilístico e que princípios de condicionalização oferecem um esquema formal de como a atualização de tais graus deve ser realizada à medida em que se ganha evidência. Mas geralmente Bayesianos subjetivos não propõem outras restrições sobre atribuições de probabilidade inicial (*priors*). Por outro lado, Bayesianos objetivos alegam que princípios de natureza *a priori* restringem probabilidades iniciais. Nessa perspectiva, haveria um modo racional e adequado de determinar objetivamente graus de probabilidade inicial, independente de evidência adicional. O princípio de indiferença e o conceito de simplicidade são candidatos a desempenhar essa tarefa.

Mais especificamente, este trabalho pretende investigar a teoria de probabilidade e a versão de Bayesianismo objetivo de Richard Swinburne. Como veremos, Swinburne tem uma tipologia de probabilidade muito peculiar e uma concepção de Bayesianismo bastante original. Antes, no capítulo de abertura, vamos tratar de alguns pressupostos de sua teoria, como a sua defesa do contrastivismo doxástico e do modelo de crença *simpliciter*. Embora Bayesianos usualmente tratem crenças como um fenômeno gradual, Swinburne mantém-se tradicional sobre esse aspecto. Em seguida, no capítulo intermediário, vamos apresentar o maquinário elementar de probabilidades e examinar a sua classificação entre tipos distintos de probabilidade. Examinaremos brevemente alguns problemas relativos à justificação epistêmica e adequação dos *grounds* de crenças em termos de probabilidade. No capítulo final e mais importante, introduziremos aspectos técnicos do teorema de Bayes e caracterizaremos as posições de Bayesianos subjetivos e objetivos. Depois, vamos nos voltar mais diretamente ao Bayesianismo objetivo de Swinburne. Ele propõe um conjunto de critérios de probabilidade lógica para a sua teoria Bayesiana: encaixe com a evidência de fundo, poder explanatório, escopo e simplicidade. Este último é certamente o mais determinante. Outras coisas sendo iguais, a alegação é de que a hipótese mais simples terá maior probabilidade inicial e, consequentemente, maior probabilidade posterior do que as suas competidoras, especialmente em circunstâncias onde tais hipóteses obedecem igualmente bem a todos os outros critérios. No final, falaremos de problemas e objeções à teoria de Swinburne e ao Bayesianismo em sentido lato.

2 Crença e Probabilidade

Se um agente S tem a atitude proposicional de crença sobre qualquer proposição p — isto é, diríamos que *S crê que p* —, então é correto dizer que S toma o conteúdo de p como verdadeiro. Assim, se Pedro crê que *vai chover hoje em Porto Alegre*, Pedro admite a proposição de que *vai chover hoje em Porto Alegre* como verdadeira, em oposição à crença na proposição de que *não vai chover hoje em Porto Alegre*. Não significa que tal estado de coisas descrito pela crença de Pedro de fato é o caso no mundo. Pedro pode estar errado sobre tal crença, ela pode ser falsa e, presumivelmente, Pedro pode ou não ter razões ou evidências para sua crença. De todo modo, se dizemos que um agente S tem a crença de que p, então assumimos que S, alegadamente um agente doxástico, mantém uma relação com o conteúdo ou objeto da sua crença, a saber, uma proposição. Tradicionalmente, a atitude de crença é considerada um estado mental de S com um determinado conteúdo; neste caso, o conteúdo de que p. Por conseguinte, assumir que um agente doxástico S tem uma crença *simpliciter* de que p significa considerar que S tem um estado mental de crença em uma proposição p em uma instância particular de tempo t.

No entanto, certas crenças podem estar em melhor situação epistêmica do que outras: porque estão baseadas em razões e evidências adequadas, porque são formadas por processos confiáveis, porque não há contra-evidências ou derrotadores para elas, entre outras condições satisfeitas. A propósito, é um *desideratum* epistêmico que tais condições sejam conducentes à verdade, isto é, que forneçam um suporte adequado para a crença, porque têm um *status* justificador, de tal maneira que ela se aproxime da verdade. Nessa esteira, grande parte da discussão em epistemologia contemporânea se concentra na análise dessas e outras condições para a justificação de crenças e, em última instância, para conhecimento proposicional.[1]

Crença e probabilidade podem ser relacionadas de vários modos. Primeiro,

[1]Não é tarefa de tal trabalho definir o que é requerido para que um agente tenha posse de conhecimento proposicional e avaliar todas as condições para justificação epistêmica. Nosso objetivo é analisar as relações entre crença e probabilidade, as condições pelas quais uma proposição torna outra provável, tipos e interpretações distintas do conceito de probabilidade, alguns aspectos pelos quais justificação e probabilidade estão conectadas e a versão de Bayesianismo objetivo proposta por Richard Swinburne.

no modelo de crença *simpliciter* ou crença tudo-ou-nada (*all-or-nothing*), podemos explicar relações de suporte de razões ou evidências em termos de probabilidades. Supondo um conjunto de crenças A de um agente S tal que $A = \{p_1, p_2, \cdots, p_n\}$, podemos conceber que p_1 oferece suporte para p_2 que, por sua vez, oferece suporte para p_3 e assim sucessivamente. Quanto mais forte os *grounds* (ou as razões) suportam uma crença, mais tais *grounds* (ou razões) tornam tal crença provável.[2] Em outras palavras, os *grounds* são mais adequados à medida em que eles tornam a crença mais provavelmente verdadeira, ou seja, assumindo um intervalo $[0, 1]$ mapeado por uma função de probabilidades, mais próximo o suporte da crença está do valor *maximum* 1, mais justificada está a crença. Aqui, claramente, estaríamos falando de graus de justificação. Segundo, pode ser apropriado explicar a relação de suporte de uma concepção internalista ou evidencialista de justificação pelo recurso de algum tipo de probabilidade epistêmica. Poder-se-ia alegar que se um conjunto de evidências $e_1, e_2, \cdots, e_n$ fornece um determinado grau de suporte χ a uma proposição h num dado instante de tempo t, supondo que $.5 < \chi \leq 1$, então, além de h ser provável, h tem *grounds* adequados em t.[3] Mas se defendermos uma posição externalista, como o confiabilismo, crenças podem ter *grounds* mais adequados à medida em que estes as tornam estatisticamente provável. Um processo de formação de crença é considerado de um tipo confiável porque a sua *ratio* de produção de crenças verdadeiras é mais alta do que a de produção de crenças falsas. É mais provável estatisticamente que esse processo de formação de crenças produza crenças verdadeiras do que falsas, uma vez que a sua *frequência relativa* de produção de crenças verdadeiras é alta. Terceiro, diferente da concepção canônica de crença *simpliciter*, o modelo de graus de crença descreve o conceito de crença em termos de gradações. Se um agente S crê com grau χ que p e com grau ε que r tal que $\chi < \varepsilon$, então S crê mais fortemente em r do que p; ou crê mais fracamente em p do que r. Tais gradações de crença

[2]Essa é uma tese importante defendida por Swinburne. A princípio, *grounds* de uma crença podem ser de um tipo doxástico, outras crenças que a suportam, ou de um tipo não-doxástico, como experiências e sensações. Tais distinções serão objeto da seção final do capítulo 3. Mostraremos alguns empregos de tipos diferentes de probabilidade no contexto internalismo-externalismo.

[3]A sugestão de que o grau de suporte probabilístico deve ser acima de .5 parece ser insuficiente para o tipo de justificação relevante para conhecimento proposicional. Talvez seja um *desideratum* que esse grau seja muito próximo ou precisamente 1. Veremos isso no final do próximo capítulo.

são geralmente modeladas com o uso do aparato do cálculo de probabilidades, uma abordagem bastante aceita entre Bayesianos subjetivos.

Este capítulo pretende apresentar e avaliar as bases teóricas da epistemologia de Richard Swinburne no que diz respeito ao conceito de crença e a sua relação com o conceito de probabilidade.[4] Por assim dizer, são os aspectos propedêuticos que constituem o *background* conceitual necessário para uma avaliação mais minuciosa das partes importantes de sua teoria. Antes de explorarmos tais relações em mais detalhes, é importante destacar algumas distinções decisivas entre os modelos de crença *simpliciter*, o adotado por Swinburne, e o de graus de crença, a concepção mais popular entre Bayesianos. Em seguida, é de bom alvitre caracterizarmos a visão de Swinburne sobre as propriedades da crença, como a sua defesa do contrastivismo e do involuntarismo doxástico; sendo a primeira de importância fundamental para Swinburne a respeito da relação entre crença e probabilidade. Trataremos, ainda que brevemente, das definições de racionalidade epistêmica e racionalidade pragmática. Antes de procedermos às partes essenciais deste trabalho, focaremos nas relações entre crença, probabilidade e ação na última seção deste capítulo.

2.1 Crença *Simpliciter* e Graus de Crença

De acordo com a definição clássica de conhecimento proposicional, crença é uma condição necessária, ainda que não suficiente, para que um agente seja possuidor de conhecimento. Necessariamente, um agente S sabe que p somente se S crê que p, ou seja, não é possível para S saber que p sem crer p; a tese de que conhecimento acarreta crença.[5]

Apesar do conceito de crença ser um truísmo para epistemólogos, existem duas concepções distintas em relação à sua natureza. Uma primeira concepção con-

[4]Uma avaliação mais completa sobre o aparato formal do cálculo de probabilidades, tipos e interpretações diferentes de probabilidade, confirmação incremental e absoluta, Bayesianismo e outros tópicos relacionados será fornecida nos capítulos posteriores.

[5]Tradicionalmente, as condições de crença, verdade e justificação, ainda que não conjuntamente suficientes, são consideradas condições necessárias para conhecimento proposicional. Ver Edmund Gettier (1963) e Vincent Hendricks (2006, cap. 2, p. 13). Uma visão distinta, onde conhecimento é o mesmo que evidência ($E = K$) e é primitivo (não-analisável em termos de crença verdadeira justificada + alguma quarta condição), pode ser encontrada em Timothy Williamson (2000).

sidera que crenças são uma questão de tudo-ou-nada. Um agente S crê ou não que p. Assim, crença é um estado mental no qual um agente se encontra total ou completamente. Não há, destarte, gradações intermediárias de força de crença ou convicção entre um mínimo e um máximo. Se Pedro crê que *há um livro de capa vermelha na mesa à sua frente*, então ele crê completamente em tal proposição. Mais precisamente, além de crer que p, um agente poderia descrer que p, o mesmo que crer que $\neg p$, ou suspender o juízo (ou duvidar) sobre p. Tais atitudes doxásticas são tomadas em sentido absoluto ou, em sentido rigoroso, são atitudes de crença, descrença e suspensão de juízo *simpliciter*. Poderíamos, ainda, representá-las formalmente, assumindo uma função $f_{St}(\,\cdot\,)$ para tais atitudes epistêmicas de S sobre proposições de um campo ou álgebra $\mathbb{F}$ ($p \in \mathbb{F}$), da seguinte maneira: $f_{St}(p) = 1$ para crença de S em p em t, $f_{St}(p) = 0$ para suspensão de juízo de S sobre p em t e $f_{St}(p) = -1$ para descrença de S em p em t. Nesse caso, tais três valores representam atitudes epistêmicas diferentes, mas sem admitir um intervalo $[-1, 1]$ no qual valores intermediários entre um mínimo e um máximo representam gradações para outras atitudes doxásticas ou para a força ou graus de crenças.[6]

Embora a concepção de crença *simpliciter* seja a mais defendida entre epistemólogos tradicionais, modelos de representação de graus de crença (*credences*) têm ganhado mais popularidade nos últimos tempos entre epistemólogos formais.[7] Em linhas gerais, modelos de graus de crença admitem que um agente tenha uma diversidade de gradações sobre diferentes proposições. Por isso, supondo um conjunto A que representa o sistema de crenças de um agente S e que $p \in A$ e $r \in A$, S crê que p com grau χ e que r com grau ε em t. Geralmente, tais graus situam-se em qualquer valor determinado ou são representados por algum subintervalo de um intervalo maior $[0, 1]$. Graus de crença têm valores precisos ou definidos (*sharp value*)

[6]Sobre lógicas epistêmica e doxástica, o trabalho de Jaakko Hintikka (1962) é certamente precursor e de grande importância. Vincent Hendricks e John Symons (2006) cobrem uma parte substancial do assunto com um excelente *overview* e, mais recentemente, Luis Rosa (*forthcoming*, *manuscript*) tem desenvolvido um modelo formal para atribuições de racionalidade.

[7]Darren Bradley (2015) e Jonathan Weisberg (2015, secs. 1 e 2) oferecem um excelente material introdutório com enfoque no uso do maquinário Bayesiano-probabilístico em epistemologia formal. Sobre teorias e modelos de graus de crença, o volume *Degrees of Belief* editado por Franz Huber e Cristoph Schmidt-Petri (*Synthese Library 242*, 2009) é sem dúvida um dos materiais mais completos sobre o assunto.

quando, por exemplo, $\chi = .5$ e $\varepsilon = .9$, isto é, graus de crença com valores exatos. Se, por outro lado, os graus de crença de S são tais que χ é todo intervalo $[a, b]$ e ε é todo intervalo $[c, d]$, admitindo que $[a, b] \subseteq [0, 1]$ e $[c, d] \subseteq [0, 1]$ — ou seja, são subconjuntos de $[0, 1]$ —, então tais graus têm valores imprecisos, difusos ou indefinidos (*unsharp value*): por exemplo, $[a, b] = [.5, .6]$ e $[c, d] = [.7, .9]$; e quanto maior o intervalo, mais difuso é o grau de crença. Portanto, S pode crer que *todo triângulo tem três lados* com grau máximo 1, crença total, ao passo que crê com grau .5 que *a próxima jogada de uma moeda será coroa*. Mas, igualmente, S pode ter um grau impreciso ou difuso $[.5, .7]$ de que *vai chover hoje à noite em Porto Alegre*.[8]

O recurso ao maquinário formal da teoria de probabilidades tem sido canonicamente o mais aceito entre as teorias de graus de crença, ou seja, a estratégia teórica de explicar *graus de crença modelados como probabilidades subjetivas*. Modelos rivais a essa concepção são a teoria Dempster-Shafer e as medidas de possibilidade ancoradas na teoria dos conjuntos difusos (*fuzzy sets*). Não é nossa tarefa aqui, todavia, avaliar a plausibilidade de tais modelos disponíveis para graus de crença, nem se é possível unificar ou não as concepções de crença *simpliciter* e de graus de crença.[9] A estratégia de Swinburne, que é o nosso objeto de investigação, não consiste nem em reduzir crença *simpliciter* a graus de crença e vice-versa[10], tampouco assumir uma teoria epistêmica que admita a atitude doxástica de crença como um fenômeno gradual. Ele está comprometido, sobretudo, em defender a noção clássica de crença

[8] Mais informações e problemas envolvendo graus de crença difusos em Adam Elga (2010) e Darren Bradley (2015, cap. 3, p. 45-48).

[9] Modelos formais de representação de graus de crença são (*i*) teoria de probabilidade ou (*ii*) função de crença Dempster-Shafer (Dempster, 1968 e Shafer, 1976) ou (*iii*) teoria de possibilidade (Dubois e Prade, 1988) baseada no conceito de conjuntos difusos (Zadeh, 1978). São modelos que apresentam esquemas teoricamente elegantes na formalização de graus de crença, embora eles tenham propriedades lógicas distintas uns dos outros. Graus de crença em *disjunctos* incompatíveis têm a propriedade aditiva de acordo com terceiro axioma de probabilidade: se p e q são mutuamente exclusivas, então $Pr(p \vee q) = Pr(p) + Pr(q)$, considerando que $Pr : \mathbb{F} \longrightarrow \mathbb{R}$, $Pr(\,\cdot\,)$ é uma função com valores entre $[0, 1]$ e tal que $p \in \mathbb{F}$ e $q \in \mathbb{F}$. Se são graus de crença modelados pela função Dempster-Shafer, então tais graus são super-aditivos. Supondo $Bel(\,\cdot\,)$ como uma função Dempster-Shafer para crenças e que p e q pertencem a uma álgebra $\mathbb{F}$ de proposições, tal que $Bel(\,\cdot\,) : \mathbb{F} \longrightarrow [0, 1]$, temos $Bel(p) + Bel(q) \leq Bel(p \vee q)$, considerando que $Bel(p \wedge q) = 0$. Se são graus de crença em termos de medidas de possibilidade, então tais graus são maxitivos e super-aditivos. Assumindo Ω como um conjunto não-vazio de possibilidades e para quaisquer conjuntos A e B tal que $A \in \Omega$ e $B \in \Omega$ e $\Pi(\,\cdot\,) : \Omega \longrightarrow [0, 1]$, $\Pi(A \cup B) = max\ \{\Pi(A), \Pi(B)\}$. Mais informações sobre outras diferenças entre tais modelos em Franz Huber (2009, p. 4-16).

[10] Um exame sobre tentativas unificacionistas entre os modelos de crença *simpliciter* ou tudo-ou-nada e de crença em gradações em David Christensen (2004, cap. 2).

simpliciter, tal como ele apresenta em *Epistemic Justification* (2001, cap. 2) e *Faith and Reason* (1981 e 2005, cap. 1), muito embora ele faça a seguinte sugestão sobre a força da crença (*strength of belief*): 'crenças podem ser fortes ou fracas'.[11] Um dos seus objetivos, na verdade, está mais próximo de oferecer uma explicação da adequação dos *grounds* e de justificação epistêmica em termos de probabilidades do que aceitar graus de crença. Em uma passagem mais adiante em *Epistemic Justification* (2001, cap. 2, p. 36-37), Swinburne assume que a força da crença constitui-se (1) em uma atitude contrastiva, S crê que p em contraste a crer que $\neg p$, e (2) tal atitude contrastiva de crença é tornada mais inteligível, e com uma análise mais refinada, se recorrermos ao uso do conceito de probabilidade, acreditar que *p é mais provável, comparativamente, do que $\neg p$*. Veremos na próxima seção, no entanto, que sérias objeções e problemas podem ser levantados contra a sua teoria epistêmica da crença.

2.2 Contrastivismo Doxástico

Um primeiro aspecto importante da teoria de Swinburne (2001, p. 34-38 e 2005, p. 4-9) sobre as propriedades de crenças é a sua defesa do contrastivismo doxástico. Em uma classe de contraste Ω com um par de contraditórios p e $\neg p$, $\Omega = \{p, \neg p\}$, se um agente S crê que p, então S acredita que p em oposição a acreditar que $\neg p$. Em uma classe finita Ω' mais ampla de proposições mutuamente exclusivas e conjuntamente exaustivas[12] $p_1, p_2, \cdots, p_n$ — *i.e.* $\Omega' = \{p_1, p_2, \cdots, p_n\}$ — se S crê que p_1, então S acredita que p_1 em oposição a acreditar em cada uma das proposições alternativas p_i tal que $i = 2$, $i = n$, mas $i \neq 1$.

Como dissemos anteriormente, a concepção de crença sustentada por Swinburne é a tradicional entre epistemólogos. Se um agente S tem a crença de que p, então S crê que p *simpliciter*. Ademais, Swinburne procura estabelecer algumas relações lógicas entre crer que p e crer que *p é provável*. À primeira vista, parece correto afirmar que enquanto a primeira é de um tipo mais simples, que de fato atribuímos a

[11]No original: 'beliefs may be strong or weak' (2001, p. 34).

[12]Uma partição ou um conjunto finito de proposições $\{p_1, p_2, \cdots, p_n\}$ ($i = 1$, $i = 2$, $\cdots$, $i = n$) no qual uma e somente uma proposição p_i é verdadeira. Mutuamente exclusivas: apenas uma pode ser verdadeira; conjuntamente exaustivas: pelo menos uma deve ser verdadeira.

agentes ordinários, a segunda é de um tipo particularmente mais sofisticado, da qual apenas agentes que têm algum conceito de probabilidade, ainda que não rigorosamente definido e não qualificado, poderiam ter. De fato, crenças sobre probabilidades parecem ser diferentes de crenças simples. Suponha um determinado mecanismo de jogadas de moeda com somente dois resultados possíveis, um conjunto com duas possibilidades ou uma partição tal que $W = \{cara, coroa\}$. Agora, para efeitos de distinção, vamos observar as duas alegações de crença seguintes: (a) S crê que *o resultado deste mecanismo de jogadas de moedas será coroa*; (b) S crê que *é provável*[13] *que o resultado deste mecanismo de jogadas de moedas será coroa*. Podemos nos perguntar: um agente S pode crer que p, mas não crer que *p é provável*? Ou seja, se é possível para S crer na primeira (a) sem crer na segunda (b); presumivelmente, de que o agente crê que o evento descrito por essa afirmação tem probabilidade alta ou maior que .5 ($\frac{1}{2}$). Com efeito, a crença descrita em (a) incide sobre uma determinada proposição p, enquanto a crença expressa em (b) incide sobre uma outra proposição: $provavelmente, p$. Mas considere a natureza absurda das seguintes asserções: 'eu creio que está ensolarado em Porto Alegre, mas é improvável que isto é o caso' e 'eu creio que está ensolarado em Porto Alegre, mas não creio que isto é provável'.

O exemplo acima poderia nos motivar a dizer que, na verdade, a primeira crença é melhor explicada em termos de crença sobre probabilidades. Por um lado, poder-se-ia alegar que crer que 'o resultado deste mecanismo de jogadas de moedas será coroa' significa o mesmo que crer que 'o resultado deste mecanismo de jogadas de moedas será provavelmente coroa', em comparação ou oposição ao resultado de *cara*. Todavia, isso envolveria a exigência de que o agente tivesse algum conhecimento ou crença de fundo (*background belief*) a respeito de alguma noção de probabilidade. Por outro lado, pode-se argumentar que embora S creia *simpliciter* que p sem crer *explicitamente* que *p é provável*, isso não exclui a possibilidade de S crer *implicitamente* que *p é provável*, no sentido de não entreter explicitamente o conteúdo da sua crença no pensamento, quando S crê *simpliciter* que p. De qualquer

[13] *Sub judice* o tipo de probabilidade envolvido nessa afirmação, se probabilidade física, propensão natural do mundo em produzir certo tipo de evento, ou estatística, frequência relativa na qual um tipo de evento ocorre ou de alguma outra natureza. Uma avaliação mais completa do assunto será oferecida no próximo capítulo.

forma, Swinburne muitas vezes sugere, especialmente na primeira edição de *Faith and Reason* (1981, p. 4-8), que crer que p equivale a crer que *p é mais provável do que $\neg p$*. Assim, uma vez que ele considera crença como uma atitude contrastiva, crer que p equivale a crer que p em oposição a crer que $\neg p$ que, por seu turno, equivale a crer que *p é mais provável do que $\neg p$*. Nesse sentido, supondo uma classe de contraste com um par de contraditórios, Swinburne (2001, p. 35) admite que *p é mais provável do que $\neg p$* quando a probabilidade de p é maior do que .5, o que significa que a probabilidade de $\neg p$ é menor do que .5. Isso é bastante correto e razoável, visto que, de acordo com um teorema demonstrável com algumas suposições e axiomas do cálculo de probabilidades, temos $Pr(\neg p) = 1 - Pr(p)$. Portanto, se $Pr(p) > .5$, então $Pr(\neg p) < .5$.[14]

No entanto, a tese de que crer que p sempre corresponde a que crer que *p é mais provável do que $\neg p$* não passa incólume a uma crítica severa. Nesse espírito, William Alston (1994, p. 25-29) oferece duas objeções à tese de Swinburne. Primeira, crer que p nem sempre equivale logicamente a crer que *p é mais provável do que não-p*. Várias das nossas crenças são crenças *simpliciter*, sem necessariamente serem crenças sobre probabilidades. Nessa perspectiva, Alston formula o seguinte contraexemplo: agentes ordinários, como crianças de pouca idade, têm crenças *simpliciter* em várias proposições, mas não é o caso que, para todas tais crenças, tais agentes igualmente acreditem que uma certa proposição é mais provável do que a sua negação ou do que alternativas incompatíveis. O ponto nevrálgico de Alston consiste no fato de que agentes menos sofisticados talvez nem tenham alguma ideia geral de probabilidade no seu *background* conceitual; isso não quer dizer, porém, que tais agentes não tenham em absoluto crenças simples. Fora a questão sobre o que é ter posse de um conceito — se articulá-lo linguisticamente é uma condição necessária ou não para tal — ainda que agentes menos sofisticados não tenham o conceito de probabilidade, tais indivíduos podem crer *como se* tivessem tal conceito ou *como se p fosse provável* para eles; talvez no sentido de crer *implicitamente* que

[14]Demonstrações de teoremas estão disponíveis no apêndice deste trabalho. Podemos similarmente formular tais relações em termos de probabilidade condicional. Para uma mesma evidência ou conjunto de evidências total e_t, h é mais provável do que $\neg h$, ambos condicionados em e_t, somente se $Pr(h \mid e_t) > Pr(\neg h \mid e_t)$. Por conseguinte, se $Pr(h \mid e_t) > .5$, então $Pr(\neg h \mid e_t) < .5$.

p é provável. Segunda, a alegação de que crer que p é crer que *p é mais provável do que* $\neg p$ conduz a um regresso vicioso *ad infinitum*. Trata-se de uma crítica mais contundente e decisiva de Alston. Como vimos, a crença de um agente S de que p é explicável em termos de S crê que *p é mais provável do que* $\neg p$. Esta última, por sua vez, também é uma crença. Assim, S crê que *p é mais provável do que* $\neg p$ é reduzida a uma outra crença de S, nomeadamente, à crença de que (*p é mais provável do que* $\neg p$) *é mais provável do que não é o caso que* (*p é mais provável do que* $\neg p$) e assim infinitamente. O problema de tal regresso vicioso é que ele se alastra em direção a todas as crenças do sistema de crenças de um agente S, tornando-as infinitamente complexas. Mais seriamente, admitindo que crer que p é sempre crer que *p é mais provável do que não-p*, o problema se estende a todas as crenças de todos agentes doxásticos. E tal consequência não é exatamente desejável, nem em termos explicativos, tampouco do ponto de vista epistêmico.

A boa notícia é que o próprio Swinburne acusa o golpe e reconhece a força das duas objeções de Alston numa nota de *Epistemic Justification* (2001, p. 36). Na edição mais atualizada de *Faith and Reason* (2005, p. 5), Swinburne afirma que: 'normalmente, crer que p é crer que *p é mais provável do que não-p*'.[15] Não sempre, mas na maioria dos casos, o que ainda é bastante implausível se tratando de agentes epistêmicos ordinários. De todo modo, Swinburne (2001, p. 35-36) enfraquece a sua tese e propõe relações lógicas entre crença *simpliciter* e crença sobre probabilidades com base em condições mais plausíveis e apropriadas. Sob as suposições de que $Pr(\,\cdot\,)$ é uma função probabilística, $\mathbb{F}$ é uma álgebra ou um campo de proposições sobre um conjunto de possibilidades Ω tal que $\Omega \neq \emptyset$, $p \in \mathbb{F}$ e $Pr : \mathbb{F} \longrightarrow [0, 1]$, temos as seguintes relações (Swinburne, 2001, p. 36):

- (1) S crer que $Pr(p) > .5$ acarreta S crer que p;
- (2) S crer que $Pr(p) < .5$ acarreta S não crer que p;
- (3) S crer que $Pr(p) = .5$ acarreta nem crença de S de que p, nem crença de S de que $\neg p$.

[15] No original: 'normally, to believe that p is to believe that p is more probable or more likely than not-p' (2005, p. 5).

Algumas considerações são cruciais. Primeira, a condição (1) parece correta, assumindo que .5 é um limiar de valor (*threshold*) apropriado[16] e $Pr(\neg p) = 1 - Pr(p)$. Não fica claro na exposição de Swinburne se S crer que $Pr(p) > .5$ acarreta S crer que $Pr(\neg p) < .5$, a saber, se crer que *é provável que p* acarreta crer que *não é provável que ¬p*.[17] Neste caso, estaríamos avaliando relações entre diferentes crenças sobre probabilidades, não entre crenças *simpliciter* e crenças sobre probabilidades. Vamos assumir que S crê que $Pr(p) > .5$. De acordo com (1), podemos inferir que S crê que p. Se S crê que p, então *não é o caso que* S crê que $\neg p$. De *não é o caso que* S crê que $\neg p$, usando (1) do sistema proposto por Swinburne, deduzimos por *modus tollens* que *não é o caso que* S crê que $Pr(\neg p) > .5$. Desse modo, S crer que $Pr(p) > .5$ acarreta que *não é o caso que* S crê que $Pr(\neg p) > .5$. Mas, plausivelmente, *não é o caso que S crê que* $Pr(\neg p) > .5$ não acarreta que S *crê que* $Pr(\neg p) < .5$. Ou seja, de S não crer que $\neg p$ *é provável* não se segue que S crê que $\neg p$ *é improvável* ou *não-provável*.[18] Por outro lado, a condição (2) não estabelece que S crer que $Pr(p) < .5$ acarrete crença de S de que $\neg p$, unicamente que não crer que p é acarretado pela crença de que *não é provável que p*. Embora a implicação seja de fato mais fraca do que S *crer que* $\neg p$, esta última não é logicamente incompatível com a consequência extraída por Swinburne. De fato, S *não crer que p* é uma asserção bem mais fraca do que S *crer que* $\neg p$. No entanto, tal como em (1), poderíamos formular (1'): S crer que $Pr(\neg p) > .5$ acarreta crença de S de que $\neg p$; crença de que $\neg p$ *é provável* acarreta crença de que $\neg p$. Quanto à condição formulada em (3), vale observar que S crer que $Pr(p) = .5$ não equivale a S duvidar ou suspender o juízo sobre p. Se a atitude epistêmica de dúvida sobre p é, por definição, equivalente a nem crer que p e nem descrer que p, então corremos o risco de confundirmos o estado mental de dúvida com o fato de um agente não ter sequer considerado p ou falhado em considerar p. Por exemplo, se Pedro falha ou deixa de considerar que

[16] Pode haver disputa se .5 é um valor arbitrário para confirmação. Veremos isso adiante.

[17] Igualmente se S crer que $Pr(p) < .5$ acarreta ou não S crer que $Pr(\neg p) > .5$.

[18] Agradecimento ao Luis Rosa por chamar atenção para esse ponto. Além disso, poderíamos nos perguntar se é racional para S crer que $Pr(\neg p) < .5$, assumindo, naturalmente, que S crê que $Pr(p) > .5$ e que é racional para S ter tal crença. Assim, se também é racional para S crer no que é acarretado pelas suas crenças racionais, então é racional para S crer que $Pr(\neg p) < .5$, admitindo um princípio de fecho apropriado para o operador de racionalidade.

Obama está trabalhando agora na Casa Branca, então Pedro nem crê e nem descrê que *Obama está trabalhando agora na Casa Branca*; nem crê que p e nem descrê que p, Pedro não formou nenhuma das duas atitudes. Contudo, não queremos dizer que isso é o mesmo que suspensão de juízo ou dúvida.

Segunda, um dos objetivos de Swinburne (2001, p. 36-37) é caracterizar crenças como atitudes contrastivas e estas em termos de probabilidade relativa. Consegue-se, assim, um amplo espectro de diferentes forças de crenças em termos comparativos: crer que *p é muito mais provável do que $\neg p$*; crer que *p é marginalmente mais provável do que $\neg p$*; crer que *p é tão provável quanto $\neg p$*; crer que *p é menos provável do que $\neg p$*. Estas são comparações muito gerais nas quais um conceito de probabilidade relativa bastante vago está sendo empregado. Mas, em última instância, a sua proposta não parece tão promissora, pois atribuir crenças sobre probabilidades a agentes doxásticos é sobremaneira não-natural e não-intuitivo.

Terceira, uma vez que crenças sobre probabilidades parecem tão não-naturais, por que não assumir a concepção de graus de crença modelados probabilisticamente? Além de mensurar a força de crenças com o recurso do aparato formal do cálculo de probabilidades, restrições de racionalidade e coerência probabilística podem ser aplicadas a graus de crença. Crenças seriam graus subjetivos que um agente tem sobre a verdade de diversas proposições. Tais graus são mais racionais e coerentes à medida em que obedecem às restrições impostas pelo maquinário de probabilidades e pelos princípios do Bayesianismo.[19] A propósito, se agentes têm graus probabilísticos de crença — ou seja, probabilidades subjetivas —, então é um *desideratum* epistêmico que restrições adicionais, como princípios de coordenação entre tipos diferentes de probabilidade, sejam satisfeitas. O custo de tal empreendimento é de se comprometer com uma epistemologia dos graus de crença, bem como avaliar a plausibilidade de um modelo unificado entre crença *simpliciter* e graus de crença. Todavia, um epistémologo tradicional, proponente do modelo de crença *simpliciter*, poderia alegar que a tese de que agentes têm graus probabilísticos de crença é tão (ou mais) não-natural quanto a concepção de crenças sobre probabilidades. Em última aná-

[19] Analogamente, restrições lógicas dedutivas, como consistência lógica e fecho dedutivo, impõem condições de racionalidade para crença tudo-ou-nada ou *simpliciter*.

lise, dificilmente agentes ordinários teriam graus de crença com valor máximo 1 em todas tautologias e com grau mínimo 0 em todas falsidades lógicas, admitindo que a norma de coerência probabilística, uma na qual graus de crença devem satisfazer o maquinário de probabilidades, é uma norma correta para *credences*.[20]

Quarta, certamente a função de probabilidades acima carece de alguma qualificação mais específica, isto é, à qual tipo e interpretação de probabilidade a função se refere. Nesse sentido, o principal objetivo dos próximos capítulos é caracterizar claramente os vários tipos, interpretações e problemas envolvendo o conceito de probabilidade. Assim como a sua tipologia de probabilidades, a versão de Bayesianismo objetivo proposta por Swinburne tem aspectos interessantes e profícuos. Precisamos, antes, avaliar outras teses igualmente importantes de sua teoria.

2.3 Outras Propriedades de Crenças

De modo geral, o conceito de crença é definido como um estado mental ao qual o agente (ou sujeito epistêmico) tem uma relação; usualmente, tal estado mental é expresso em formato de uma atitude proposicional: S crê que p. Como dissemos acima, proposições, conteúdos significativos de sentenças declarativas, são tomadas como o objeto das crenças de agentes. Acontece que existem alguns tipos de estados mentais cujo objeto não é uma proposição; não está sob a forma de uma atitude proposicional, como o estado mental de dor. Se um agente S sente uma dor de cabeça intensa, então diríamos que ele tem uma experiência, ou está experienciando, uma determinada sensação, a saber, a sensação de uma dor de cabeça com grande intensidade. Além da natureza distinta entre tais estados mentais, um proposicional e o outro não-proposicional, Swinburne (2001, p. 38) sugere uma distinção entre estados mentais conscientes (*conscious mental states*) e estados mentais contínuos ou permanentes (*continuing mental states*). Um estado mental de dor é considerado um

[20] A tese Lockeana da crença conecta os modelos de graus de crença e crença *simpliciter*. Em linhas gerais, é racional para um agente S crer que p se e somente se o grau de crença de S em p é maior do que χ. O problema é que versões do paradoxo da loteria podem ser formulados de tal maneira que seria racional para S crer que p e crer que $\neg p$ mesmo se $\chi = 0.99$. Pode-se estancar o paradoxo da loteria se $\chi = 1$. De acordo com tal proposta, é racional para S crer que p se e somente se o seu grau de crença em p é 1, o que parece implausível para muitas das crenças que um agente tem. Mais informações sobre essa discussão em Richard Foley (2009) e James Hawthorne (2009).

estado mental consciente. Quando temos tais estados mentais, percebemos conscientemente a sua ocorrência na nossa vida mental. Diferentemente, crenças são estados mentais contínuos porque nem sempre o agente está consciente do seu conteúdo ou objeto, mas tem a capacidade de armazená-lo em sua memória e eventualmente trazê-lo à sua consciência. Em sentido estrito, não estamos necessariamente conscientes de tais estados contínuos ou permanentes a todo tempo.

Ademais, para Swinburne (2001, p. 39-40 e 2005, p. 24-26), crença é um estado mental passivo involuntário. Normalmente, crenças não estão sob o nosso controle ou vontade,[21] como muitas das nossas decisões e ações nos parecem estar. A propósito, crer em qualquer proposição não é o mesmo que realizar uma ação na concepção de Swinburne. Se um estado de coisas do mundo causa agora a crença em S de que *há um livro vermelho sobre a mesa à sua frente* — supondo que o aparato perceptual de S funciona corretamente, S responde de maneira adequada ao ambiente externo, entre outras condições —, então S não pode simplesmente escolher não crer em tal proposição pela determinação da sua vontade ou ao seu bel-prazer. Se S pudesse escolher quais crenças ele tem, então pareceria que as suas crenças são, em primeiro lugar, objetos de sua preferência e controle. Entretanto, S pode procurar por mais razões para sua crença, investigar se o conjunto total de evidências disponíveis é bom o suficiente para torná-la provável, avaliar a confiabilidade do seu processo de formação de crenças, mais outras práticas e procedimentos adequados. Nessa perspectiva, S tem, como bem reconhece Swinburne (2001, p. 40 e 2005, p. 26), algum controle dos seus métodos e práticas cognitivas de investigação e formação de crenças através do tempo, a saber, em uma avaliação diacrônica, de um tempo t para um tempo t' e assim sucessivamente.

Finalmente, Swinburne (2001, p. 38-39) defende que crenças possuem mais duas propriedades, acessibilidade privilegiada e infalível. Crenças são estados mentais cujo possuidor, o agente, tem acesso privilegiado. Os meios pelos quais S tem

[21] Contrariamente, a atitude proposicional de aceitação, S aceita que p, envolve algum tipo de controle mais direto por parte do agente. Pode haver casos em que S crê que p, mas não aceita que p; e casos onde S aceita que p, mas não crê que p. Aceitação e crença são comumente distinguidas na literatura em epistemologia. Mais informações sobre diferenças entre crença e aceitação em Jonathan Cohen (1992) e Eric Schwitzgebel (2015, sec. 2.5).

acesso direto às suas crenças não estão disponíveis aos outros. Podemos ter acesso público às crenças de S por meio dos seus atos de asserção e do seu comportamento, mas não no mesmo sentido que S tem acesso às suas próprias crenças e, sob um enfoque mais geral, aos seus próprios estados e eventos mentais.[22] Outra característica importante é o acesso infalível que agentes têm em relação às suas próprias crenças. Embora Swinburne (2001, p. 38-40) seja bastante breve sobre tal propriedade, uma primeira interpretação do que ele pretende defender com acessibilidade infalível pode ser formulada nos seguintes termos. Se S crê que p em t, então S não pode estar errado de que ele tem tal crença em t, ou seja, a crença de que p. O conteúdo proposicional da crença de S pode ser falso, talvez a acurácia da crença de S não esteja calibrada, talvez a sua crença não represente corretamente um determinado estado de coisas do mundo em t, talvez S não tenha justificação para crer que p em t. Contudo, o fato de que S tem a crença de que p em t é infalível porque caso ele creia que p em t, presumivelmente essa é uma das suas crenças em t. Seria incorreto se Swinburne quisesse dizer que infalibilidade é uma propriedade intrínseca do conteúdo da maioria de nossas crenças: muitas delas ainda podem ser falsas! Em outro momento, no entanto, Swinburne está disposto a aceitar que crenças de aparência, ou de como o mundo nos parece ser, são infalíveis, sendo esta uma interpretação mais forte sobre acessibilidade infalível. Assim, a crença de S de que *parece que* p em t é infalível no sentido de que S não pode estar errado sobre o seu conteúdo em t. S não pode estar errado de como as coisas *parecem* a ele em um dado tempo.[23]

2.4 Racionalidades Epistêmica e Pragmática

Podemos discriminar entre duas maneiras pelas quais é racional para um agente ter uma crença: uma epistêmica e outra pragmática[24] (ou prática). Quanto à

[22]Em *Mind, Brain, and Free Will* (2013), Swinburne distingue estados e eventos mentais de estados e eventos cerebrais: os primeiros são de acesso privilegiado pelo próprio agente que os possui, perspectiva de primeira pessoa, e os últimos são de acesso público a todos, perspectiva de terceira pessoa.

[23]'Since somenone's belief about something just is the way that a certain aspect of the world looks to him at a particular time, he cannot at that time be in error about it—that is, about the content of his belief' (2001, p. 39).

[24]Salvo melhor juízo, racionalidade epistêmica é distinta de racionalidade pragmática. Seguimos essa distinção junto com Thomas Kelly (2003) e David Christensen (2004, cap. 1, p. 4-5).

primeira, é racional para S crer que h com base em um conjunto total de evidências e_t disponível para S somente se e_t torna h provável. De modo converso, não seria racional para S crer que h com base em e_t em virtude de e_t tornar h improvável. Nesse primeiro sentido, dizemos que é racional *epistemicamente* para S ter a crença de que h porque h é suportada por bons *grounds* ou evidências. Quanto à segunda, é racional para S crer que h', ou agir *como se* crendo em h', porque é instrumentalmente mais vantajoso para S ter essa crença. Nesse segundo sentido, dizemos que é racional *pragmaticamente* para S ter a crença de que h', ou agir *como se* h' é verdadeira, uma vez que isso contribui para a realização de um determinado propósito de S. Em outras palavras, pode ser que crer ou agir *como se* h' é verdadeira tenha alto valor pragmático, representado em termos de utilidades, para S; como a satisfação de um dos propósitos últimos ou preferências mais fortes de S.

Nem sempre, todavia, racionalidades epistêmica e pragmática coincidem. Por um lado, supondo que as evidências disponíveis para S tornam a hipótese de que *Deus existe* improvável, não seria racional epistemicamente para S crer que *Deus existe*. Por outro lado, considerando que agir *como se* crendo que *Deus existe* é vantajoso para S — a utilidade de crer que *Deus existe* é maior do que a de não crer que *Deus existe*, ela traz ganhos práticos para S, como uma boa vida e a bem-aventurança, e evita *des*utilidade, como a miséria e a infelicidade —, seria racional pragmaticamente para S agir *como se* a proposição *Deus existe* fosse verdadeira.[25]

A distinção entre os dois tipos de racionalidade é equivalentemente caracterizada pela diferença dos fins ou objetivos envolvidos. Tipicamente, fins epistêmicos são crer em verdades e não crer em falsidades, ao passo que fins pragmáticos ou práticos dizem respeito a agir *como se* tal e tal é verdadeiro à medida em que isso produz algum benefício prático do qual o agente estima grande valor para si; ou, similarmente, racionalidade epistêmica refere-se a razões que são conducentes ou aproximam a crença da verdade e racionalidade pragmática diz respeito a quais

[25] Trata-se de uma versão simplificada da aposta de Pascal, com intuito de distinguir os dois sentidos de racionalidade. Nesse caso, sob a suposição de que h = *Deus existe*, temos quatro alternativas: ganho, crer que h e h é verdadeira; perda, crer que $\neg h$ e h é verdadeira; nem ganho e nem perda, crer que h e $\neg h$ é verdadeira; nem ganho e nem perda, crer que $\neg h$ e $\neg h$ é verdadeira. Mais detalhes e diferentes versões da aposta de Pascal em Ian Hacking (2001, cap. 10, p. 114-126) e Alan Hájek (2012).

são os meios que maximizam um propósito pretendido pelo agente. A despeito de tais diferenças, muitas das nossas crenças têm impacto sobre as nossas ações. Em particular, podemos ter crenças sobre qual curso de ação adotar porque provavelmente ele nos conduzirá a alcançar um determinado objetivo ao qual conferimos alto valor prático. Vamos investigar na próxima seção como os conceitos de crença, probabilidade e ação estão conectados de acordo com a proposta de Swinburne.

2.5 Crença, Probabilidade e Ação

Suponha que S tenha o propósito de chegar de ônibus ao *campus* central da Universidade onde ele estuda. Para tanto, suponha que S tenha à sua disposição somente duas ações A_1 e A_2, isto é, pegar a linha de ônibus $T1$ ou $T2$. Ademais, assumindo outras crenças de fundo (*background beliefs*) de S — a crença de que as duas linhas chegam no mesmo horário na parada de ônibus onde S está, de que o preço do bilhete é o mesmo, entre outras —, se S crê que é mais provável que A_1 alcance o seu propósito pretendido do que A_2, então se S é racional, ele realizará a ação A_1, pegar a linha de ônibus $T1$ em vez de pegar a linha $T2$. Neste caso, S tem uma crença sobre qual curso de ação provavelmente levará à consecução do seu objetivo de chegar ao *campus* central da sua Universidade, considerando que esse é o único ou o mais forte dos propósitos de S.

Nessa perspectiva, Swinburne (2001, p. 42 e 2005, p. 10) caracteriza tal crença de S como uma crença de um tipo específico: *crenças de meios a fins*.[26] Como mencionamos anteriormente, são crenças sobre qual ação de uma partição finita de ações básicas $A_1, A_2, \cdots, A_n$ provavelmente realizará algum dos objetivos incompatíveis $O_1, O_2, \cdots, O_n$ pretendidos por S; por exemplo, a crença de que é mais provável que A_1 atingirá O_1 do que qualquer outra ação alternativa A_i, dado que $i \neq 1$. Ações básicas ou de curto-prazo são, para Swinburne (2001, p. 42-44), ações que produzem outras ações, de médio ou longo prazo, que, por sua vez, levam à realização de um objetivo ou fim. Elas são formadas por propósitos mais básicos

[26]'S's beliefs about the probability of different actions attaining his goal are what I shall call means-ends beliefs' (2005, p. 10).

de S e podem gerar uma longa cadeia causal de ações; estas últimas correspondem a propósitos de médio ou longo prazo. Além disso, tal cadeia causal de ações resulta em algum propósito último de S, que não é uma parte intermediária ou de transição entre uma ação mais básica e um propósito de longo prazo, a saber, não há nenhum outro propósito de longo prazo ao qual um propósito último está subordinado.[27] Por exemplo, S acredita que uma ação, a de andar por uma determinada rua, provavelmente realizará um propósito intermediário, de chegar a um ponto de ônibus, que, por seu turno, provavelmente realizará o propósito último de S, que corresponde a ir de ônibus ao *campus* central da Universidade onde ele estuda.

Destarte, Swinburne (2001, p. 44) nos oferece uma análise de como crenças de meios a fins têm consequências importantes para ações e propósitos últimos. Suponha uma partição finita Ω de propósitos últimos de S tal que $\Omega = \{O_1, O_2, \cdots, O_n\}$; cada um dos propósitos corresponde a um determinado objetivo ou fim pretendido por S. Igualmente, suponha uma partição finita de ações W de S tal que $W = \{A_1, A_2, \cdots, A_n\}$; cada ação corresponde a diferentes propósitos mais básicos e S realizaria uma das ações somente se formasse o propósito mais básico correspondente. Em um primeiro cenário, O_1 é o propósito último mais forte de S tal que $\Omega' \subset \Omega$ e, alegadamente, $\Omega' = \{O_1\}$; Ω' é um subconjunto de Ω com apenas um propósito último de S, o mais forte deles. Assim, S formará um propósito mais básico cujo resultado é A_1 porque S crê que é mais provável que A_1 realize O_1 do que qualquer ação alternativa A_i, dado que $A_i \in W$ e $i \neq 1$. Em outros termos, de cada ação possível disponível, S crê é mais provável que A_1 alcance O_1 do que qualquer outra ação: S crê que é mais provável que A_1 alcance O_1 do que A_2, S crê que é mais provável que A_1 alcance O_1 do que A_3 e assim por diante. Em contrapartida, um segundo cenário descreve O_1 como um propósito tão forte quanto outros propósitos últimos de S. Suponha, agora, que Ω'' represente os propósitos mais fortes de S tal que $\Omega'' \subset \Omega$ e $\Omega'' = \{O_1, O_2, O_3\}$; um subconjunto de propósitos últimos incompatíveis. Presumivelmente, S formará um propósito mais básico cujo resultado é A_i

[27]'A purpose that is not explained by any longer-term purpose I will call an ultimate purpose. The agent does not seek to fulfil an ultimate purpose as a step towards any further goal' (2001, p. 43).

porque S crê que é mais provável que uma ação A_i de W realizará qualquer um dos propósitos de Ω'' do que qualquer ação alternativa A_k de W, dado que $i \neq k$.

Crenças de meios a fins, no entanto, pressupõem outras crenças. Nesse sentido, S crer que a linha de ônibus $T1$ provavelmente vai satisfazer o seu propósito de chegar à Universidade envolve uma conjunção de outras crenças do sistema doxástico de S: a crença de que as linhas de ônibus estão funcionando normalmente na sua cidade, a crença de que a linha de ônibus $T1$ passará em breve no ponto de ônibus onde S está, *etc.* Mais precisamente, Swinburne (2001, p. 45) reconhece que (1) crenças de meios a fins podem ser o resultado de diversos conjuntos de crenças teóricas, que representam partes de diferentes sistemas doxásticos que S poderia ter, (2) tais crenças teóricas são entendidas como crenças sobre como o mundo parece ser ou é categoricamente para S e (3) qualquer ação A_i produzida por um propósito mais básico, e que visa qualquer propósito último e forte O_i de S, é compatível com vários conjuntos alternativos de crenças teóricas que S poderia ter.[28]

Por fim, podemos fazer uma observação crítica interessante sobre a proposta de Swinburne. Dissemos que S realizará A_i porque S crê que é mais provável que uma ação A_i alcançará o propósito mais forte de S, ou qualquer um dos mais fortes, do que qualquer ação alternativa A_k, $i \neq k$. Embora Swinburne proceda a uma avaliação em termos de probabilidade comparativa sem ainda oferecer uma tipologia mais precisa, não é claro se o único requerimento necessário para que uma ação A_i seja

[28]Uma maneira alternativa de conectar crenças e consequências das ações é pelos teoremas representacionais e princípio de máxima utilidade esperada em teoria da decisão. Assim, para uma determinada ação A e uma partição de consequências $\{C_1, C_2, \cdots, C_n\}$, mede-se em uma função única a utilidade esperada da escolha pela ação A, isto é, pela soma dos produtos de graus probabilísticos de crença e utilidades para cada uma das consequências da partição: $Exp_{Pr,U}(A) = \sum[Pr(C_i \mid A) \times U(C_i)]$. Além disso, escolhas por uma ação A, B, C, entre outras, precisam satisfazer um conjunto de restrições de racionalidade. Em resumo, teoremas representacionais pretendem demonstrar que se o agente satisfaz um determinado conjunto de restrições ou condições de racionalidade, então as suas preferências e probabilidades subjetivas podem ser representadas como resultantes dessa função única de utilidade esperada. Um dos problemas dessa proposta é a violação da restrição de independência e o paradoxo de Allais; por exemplo, em certos cenários agentes escolhem pela maior utilidade esperada e em outros cenários escolhem pela menor utilidade esperada ou são pró-risco em certas circunstâncias e têm aversão ao risco em outras. De qualquer forma, em teoria da decisão, a proposta de Leonard Savage (1954) é tomada como o modelo *standard* entre Bayesianos contemporâneos — por exemplo, Patrick Maher (1993) e James Joyce (2004) — e é considerada uma teoria normativa sobre escolhas e decisões em situações de incerteza. Um *overview* sobre teoria da decisão, teoremas representacionais, problemas e propostas está disponível em K. Steele e H. O. Stefánsson (2015).

mais provável do que uma outra alternativa A_k, no intuito de atingir um propósito último, limita-se unicamente ao fato de que o grau de probabilidade de A_i alcançar tal propósito é maior do que o grau de A_k alcançá-lo. Ou se, além disso, é exigido que o grau de probabilidade de A_i alcançar o propósito último pretendido por S seja superior a um limiar de valor específico χ, um no qual, tipicamente, $\chi = .5$. Em todo caso, como reiteramos em algumas passagens deste capítulo, estamos falando de crenças sobre probabilidades, algo que, intuitivamente, atribuímos somente a agentes doxásticos sofisticados.

3 Probabilidade

Até aqui tratamos de um conceito vago e pouco preciso de probabilidade relativa ou comparativa e, especialmente, falamos da sua relação com a concepção de crença *simpliciter*. Neste capítulo, entretanto, vamos definir e discutir vários usos e diferentes conceitos de probabilidade de acordo com a classificação proposta por Swinburne: probabilidade física ou natural (*propensity-type*), probabilidade estatística (*frequency-type*) e tipos distintos de probabilidade indutiva (subjetiva, epistêmica e lógica). Antes de introduzir e avaliar a tipologia de Swinburne, apresentaremos o aparato formal básico do cálculo de probabilidades e algumas definições centrais; exceto o teorema de Bayes, objeto do capítulo 4. Após uma análise cuidadosa dos diversos tipos de probabilidade e uma breve avaliação de dois princípios de coordenação probabilística, mostraremos que teorias distintas de justificação epistêmica, internalista e externalista, empregam conceitos divergentes de probabilidade para a adequação dos *grounds*. Em particular, uma crença tem *grounds* mais adequados, e é melhor justificada epistemicamente, somente se tais *grounds* a tornam provável — com grau χ tal que $1 \geq \chi > .5$ e, em termos de melhor otimização, quando χ se aproxima de 1, isto é, $\chi \approx 1$.

3.1 Axiomas e Definições do Cálculo de Probabilidades

Os axiomas tradicionais da teoria elementar de probabilidade[29] são conhecidos como *axiomas de Kolmogorov*.[30] Assim, assumindo que $Pr(\ \cdot\ )$ é uma função de probabilidades, Ω é um conjunto não-vazio de possibilidades — ou seja, $\Omega \neq \emptyset$ — e $\mathbb{F}$ é uma álgebra ou campo de proposições sobre Ω, temos $Pr : \mathbb{F} \longrightarrow \mathbb{R}$, sob a suposição de que $p \in \mathbb{F}$ e $q \in \mathbb{F}$, *se e somente se*:

[29] O formalismo aqui apresentado pode ser encontrado em muitos livros e artigos sobre probabilidade e Bayesianismo: Paul Horwich (1982), John Earman (1992), Ian Hacking (2001), Richard Jeffrey (2004), James Joyce (2004 e 2011), Howson e Urbach (2006), Franz Huber (2009), Alan Hájek (2011), Jonathan Weisberg (2011 e 2015) e Michael Strevens (*manuscript*).

[30] Ou axiomas do *calculus* de probabilidades. Na formulação clássica, os axiomas são construídos com a linguagem e os operadores da teoria dos conjuntos: (1') $Pr(A) \geq 0$, (2') $Pr(\Omega) = 1$ e (3') Se $A \cap B = \emptyset$, $Pr(A \cup B) = Pr(A) + Pr(B)$ e $Pr(A \cap B) = 0$, dado que $\Omega \neq \emptyset$, $\mathbb{F}$ é um conjunto de subconjuntos de Ω que contém Ω como membro, $A \in \mathbb{F}$, $B \in \mathbb{F}$ e $Pr : \mathbb{F} \longrightarrow \mathbb{R}$. Disponível no clássico *Foundations of the Theory of Probability* de A. N. Kolmogorov (1956 [1933], cap. 1, p. 2).

- (1) $Pr(p) \geq 0$;
- (2) Se p é uma tautologia, então $Pr(p) = 1$;
- (3) $Pr(p \vee q) = Pr(p) + Pr(q)$ se p e q são mutuamente exclusivas (ou incompatíveis).[31]

Os axiomas (1), (2) e (3) são conhecidos, respectivamente, como o de não-negatividade, o de normalização (ou certeza) e o de aditividade finita. Ademais, os três componentes Ω, $\mathbb{F}$ e Pr formam o que se entende por espaço de probabilidade (*probability space*): $\langle \Omega, \mathbb{F}, Pr \rangle$.[32] Em tal estrutura formal, $Pr(\,\cdot\,)$ é uma função sobre $\mathbb{F}$ e os valores de probabilidade atribuíveis aos *relata* dessa função estão encerrados em um intervalo $[0, 1]$, incluindo os valores respectivos aos elementos do conjunto dos reais $\mathbb{R}$. Em outras palavras, $Pr(\,\cdot\,)$ mapeia os elementos de $\mathbb{F}$ ao conjunto $\mathbb{R}$ no intervalo $[0, 1]$. Alegadamente, assumiremos que proposições são os *relata* ou objetos da função de probabilidade e que $\mathbb{F}$ é um *powerset* de Ω ou uma σ-álgebra (σ-*field*). A princípio, o axioma (2) inclui exclusivamente verdade lógicas (ou tautologias) da lógica clássica e $\mathbb{F}$ obedece a tal restrição.[33]

Algumas consequências importantes podem ser deduzidas dos axiomas de probabilidade.[34] Pelo axioma (2) e supondo que $\mathbb{F}$ satisfaz a lógica clássica, $(p \vee \neg p)$ e $\neg(p \wedge \neg p)$ têm o valor máximo de 1, ou seja, $Pr(p \vee \neg p) = 1$ e $Pr[\neg(p \wedge \neg p)] = 1$. Assim, pelo axioma (3), $Pr(p \vee \neg p) = Pr(p) + Pr(\neg p)$ porque p e $\neg p$ são mutuamente exclusivas. Uma vez que $Pr(p \vee \neg p) = 1$, conseguimos, por substituição, $Pr(p) + Pr(\neg p) = 1$. Portanto, segue-se que $Pr(\neg p) = 1 - Pr(p)$ ou $Pr(p) = 1 - Pr(\neg p)$. Por seu turno, falsidades da lógica clássica têm valor mínimo 0. Considere, por exemplo, a contradição $(p \wedge \neg p)$. Sob as suposições de que $Pr[\neg(p \wedge \neg p)] = 1$ e $Pr(p \wedge \neg p) = 1 - Pr[\neg(p \wedge \neg p)]$, infere-se que $Pr(p \wedge \neg p) = 0$.

Um teorema muito útil do maquinário de probabilidades diz que se p acar-

[31] Uma versão alternativa do axioma (3) é conhecida como aditividade contável (*countable additivity*). Em tal cenário, $\mathbb{F}$ é um conjunto fechado sob infinitas disjunções countáveis e $Pr(p_1 \vee p_2 \vee p_3 \vee \ldots\ p_n) = Pr(p_1) + Pr(p_2) + Pr(p_3) + \cdots + Pr(p_n)$, pois o conjunto $\{p_1, p_2, p_3, \ \ldots\ , p_n\}$ forma uma partição infinita de proposições.

[32] Alternativamente, o conjunto Ω é chamado de *espaço de possibilidades* ou *espaço de amostra*.

[33] A relação de acarretamento ou consequência lógica assumida aqui também é da lógica clássica. Nesse sentido, o axioma (2) poderia ser formulado de modo diferente: se $\vDash a$, então $Pr(a) = 1$.

[34] Ver apêndice para provas mais completas desses e outros teoremas.

reta[35] q, então a probabilidade de p é menor ou igual a de q. Em outros termos, se $p \vDash q$, então $Pr(p) \leq Pr(q)$. Por exemplo, para qualquer x, se *x é vermelho* (p) acarreta *x é colorido* (q), então a probabilidade de *x ser vermelho* não pode ser maior do que a probabilidade de *x ser simplesmente colorido*, a saber, $Pr(p) \leq Pr(q)$. Além disso, de acordo com o teorema 5 do apêndice, se $p \vDash q$ e $Pr(\neg p \wedge q) = 0$, então $Pr(p) = Pr(q)$. Mas se $p \vDash q$ e $Pr(\neg p \wedge q) > 0$, então $Pr(p) < Pr(q)$.

Ainda precisamos demonstrar que o valor máximo da função é 1; observe que o axioma (1) estabelece somente um valor mínimo, mas não o topo do intervalo. Para qualquer proposição p de $\mathbb{F}$, $Pr(p) = \varepsilon$ e, pelo axioma (1), $\varepsilon \geq 0$. Se p acarreta logicamente uma tautologia a, tal que $a \in \mathbb{F}$, então $Pr(p) \leq Pr(a)$. Ora, $Pr(a) = 1$ pelo axioma (2). Portanto, $Pr(p) \leq 1$. Podemos dizer, consequentemente, que para qualquer p, $Pr(p) = \varepsilon$ e $\varepsilon \in [0, 1]$ ou $0 \leq Pr(p) \leq 1$.

O axioma (3) estabelece que a probabilidade da disjunção entre duas proposições mutuamente exclusivas p e q é o resultado da adição da probabilidade de p com a probabilidade de q. Acontece que, neste caso, $Pr(p \wedge q) = 0$ porque tão somente uma delas pode ser verdadeira, ambas não podem ser verdadeiras conjuntamente. Todavia, existem situações nas quais duas proposições não são mutuamente exclusivas e a probabilidade da conjunção entre elas pode ter valor diferente de 0. Por isso, temos a seguinte formulação geral do teorema conhecido como *overlap*:

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

Como corolário óbvio da fórmula acima, segue-se que:

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

Se duas proposições são logicamente equivalentes, então os seus valores de probabilidade são iguais. Por conseguinte, $Pr(p) = Pr[(p \wedge q) \vee (p \wedge \neg q)]$ porque $p \equiv [(p \wedge q) \vee (p \wedge \neg q)]$. Considere o exemplo dos corvos pretos. *Todos os corvos*

[35] Se p acarreta semanticamente q, $p \vDash q$, então em todos mundos possíveis (ou modelos) onde p é verdadeira, q também é verdadeira. Ou, de outra forma, $p \vDash q$ se e somente se é impossível logicamente q ser falsa quando p é verdadeira.

são pretos é equivalente, por transposição, a *todas as coisas não-pretas são não-corvos*. Desse modo, uma vez que $[\forall x(Rx \rightarrow Bx)] \equiv [\forall x(\neg Bx \rightarrow \neg Rx)]$, temos $Pr[\forall x(Rx \rightarrow Bx)] = Pr[\forall x(\neg Bx \rightarrow \neg Rx)]$.

A função de probabilidade acima diz respeito à definição de probabilidade categórica ou incondicionada. Por exemplo, suponha um mecanismo de jogadas de moeda justo (*fair*) com apenas dois resultados possíveis: $\Omega = \{cara, coroa\}$. Além da moeda ser fisicamente simétrica, significa que os resultados do mecanismo de jogadas em diferentes instâncias $t_1, t_2, \cdots, t_n$ não se influenciam entre si, são independentes, e o mecanismo de jogadas não é enviesado, ou não é viciado, em um determinado resultado possível. *Ceteris paribus*, parece correto dizer que a probabilidade de cada um dos resultados possíveis de Ω é $\frac{1}{n}$, sendo n o número de possibilidades do espaço de amostra ou de possibilidades do conjunto Ω. Portanto, temos $Pr(\textit{obter cara}) = \frac{1}{2}$ e $Pr(\textit{obter coroa}) = \frac{1}{2}$. Embora seja outro exemplo bastante recorrente, considere um *dado* de seis faces e um mecanismo justo de jogadas de tal *dado*. Agora, o espaço de possibilidades é constituído de seis resultados possíveis, a saber, $\Omega' = \{1, 2, 3, 4, 5, 6\}$. Da mesma forma, assumindo as condições anteriores, diríamos que a probabilidade de qualquer um dos resultados de Ω' é ou tende a ser $\frac{1}{n}$; neste caso, $\frac{1}{6}$. Analogamente, a probabilidade de obtermos um número ímpar é $\frac{1}{2}$ porque somente três resultados são ímpares dos seis possíveis de Ω' e o resultado é um número par ou um número ímpar; ou, pelo axioma (3), $Pr(\textit{obter } 1 \vee \textit{obter } 3 \vee \textit{obter } 5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$, uma vez que os resultados de 1, 3 e 5 não podem ocorrer conjuntamente. A despeito da trivialidade dos exemplos, típicos de contextos de jogos e apostas, eles descrevem situações nas quais a definição de probabilidade categórica se aplica.

Nesta altura, podemos traçar uma distinção entre as definições de probabilidade categórica e probabilidade condicional. Considerando o mesmo mecanismo justo de jogadas de um *dado* do exemplo acima, vamos determinar a probabilidade de obtermos o número 6 como resultado condicional no fato de que se trata de um número par. Dessa maneira, o subconjunto $\Omega'' = \{2, 4, 6\}$ de Ω' abrange as três possibilidades de obtermos o resultado de número par. Portanto, pelo axioma (3), segue-se que $Pr(\textit{obter } 2 \vee \textit{obter } 4 \vee \textit{obter } 6) = \frac{1}{2}$, uma vez que os três resultados são incompatíveis. No entanto, pretendemos determinar a probabilidade de obter 6 como resultado, assumindo que o resultado é par, isto é, $Pr(obter\ 6 \mid obter\ par)$. Das três possibilidades de resultado de número par, somente uma delas é o resultado de número 6. Por conseguinte, $Pr(obter\ 6 \mid obter\ par) = \frac{1}{3}$ porque, restringindo o nosso espaço de possibilidades a três resultados, somente o número 6 satisfaz as duas condições seguintes: ser propriamente o número 6 e ser par.

Supondo que $Pr(\cdot \mid \cdot)$ é uma função de probabilidade condicional, $Pr(q) > 0$ (a função é indefinida[36] se $Pr(q) = 0$) e $Pr(\cdot \mid \cdot) : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{R}$, pode-se formalmente definir probabilidade condicional, para quaisquer p e q de $\mathbb{F}$, do seguinte modo:

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

A probabilidade de p condicional em q, $Pr(p \mid q)$, significa o mesmo que a probabilidade de p, dado que q é o caso; ou o quão provável é p sob a suposição de que q. Uma consequência trivial de tal definição é $Pr(p \wedge q) = Pr(p \mid q) \times Pr(q)$. Como veremos, essa definição permite uma análise frutífera dos diferentes tipos de probabilidade indutiva — o quão bem uma evidência ou um conjunto de evidências suporta uma hipótese tornando-a provável — e é considerada uma peça fundamental do Bayesianismo.

Um teorema igualmente importante é o de probabilidade total. Ambos, teorema de probabilidade total e a definição de probabilidade condicional, oferecem as condições para a demonstração do teorema de Bayes. Para quaisquer proposições p e q e admitindo que $1 > Pr(p) > 0$:

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

Se, alternativamente, temos uma partição $\Omega' = \{p_1, p_2, p_3, \cdots, p_n\}$, então:

$$Pr(q) = \sum_n [Pr(p_n) \times Pr(q \mid p_n)]$$

[36]Existem modelos, como os de Rényi (1970) e Popper (2002 [1959]), que permitem que a função de probabilidade condicional seja definida mesmo quando $Pr(q) = 0$.

Outro teorema básico diz que a probabilidade de uma conjunção não pode ser maior do que a probabilidade de qualquer um dos seus *conjunctos*. Formalmente, $Pr(p \wedge q) \leq Pr(p)$ e $Pr(p \wedge q) \leq Pr(q)$. Destarte, a probabilidade da proposição *Maria é enfermeira e ativista* não pode ser maior do que a probabilidade do *conjuncto Maria é enfermeira*, tampouco maior do que a probabilidade do *conjuncto Maria é ativista*. Porém, em circunstâncias especiais nas quais p acarreta logicamente q, a probabilidade de p é igual à probabilidade da conjunção $(p \wedge q)$. Em outros termos, se $p \vDash q$, então $p \equiv (p \wedge q)$ e, portanto, $Pr(p) = Pr(p \wedge q)$.

Além disso, a definição de independência pode ser expressa como se segue. Considerando que $Pr(p) > 0$ e $Pr(q) > 0$, duas proposições p e q são independentes *se e somente se* $Pr(p \mid q) = Pr(p)$. Dois apontamentos merecem o nosso destaque. Primeiro, se combinamos independência e probabilidade condicional, obtemos um resultado interessante. Suponhamos que $Pr(p) > 0$ e $Pr(q) > 0$ e que p e q são independentes entre si. Portanto, $Pr(p \mid q) = Pr(p)$ pela definição de independência. Pela definição de probabilidade condicional, temos $Pr(p \mid q) = \frac{Pr(p \wedge q)}{Pr(q)}$. Substituindo $Pr(p \mid q)$ por $Pr(p)$ na última fórmula, depreendemos que $Pr(p) = \frac{Pr(p \wedge q)}{Pr(q)}$. Logo, $Pr(p \wedge q) = Pr(p) \times Pr(q)$. Segundo, independência é uma relação simétrica. Significa que se $Pr(p \mid q) = Pr(p)$, então $Pr(q \mid p) = Pr(q)$ e se $Pr(q \mid p) = Pr(q)$, então $Pr(p \mid q) = Pr(p)$. Logo, $Pr(p \mid q) = Pr(p)$ *se e somente se* $Pr(q \mid p) = Pr(q)$. Se p e q são independentes, então p é irrelevante ou não contribui para a probabilidade de q e vice-versa.

Por último, se p e q são correlacionadas positivamente, então $Pr(p \wedge q) > Pr(p) \times Pr(q)$. Assim, segue-se que $\frac{Pr(p \wedge q)}{Pr(q)} > Pr(p)$. Portanto, pela definição de probabilidade condicional, $Pr(p \mid q) > Pr(p)$; pela mesma razão, $Pr(q \mid p) > Pr(q)$. Se p e q são correlacionadas negativamente, então $Pr(p \wedge q) < Pr(p) \times Pr(q)$. Como corolário, segue-se que $Pr(p \mid q) < Pr(p)$; equivalentemente, $Pr(q \mid p) < Pr(q)$.

3.2 Tipos de Probabilidade

É muito comum nos depararmos com alegações sobre probabilidade em diversos contextos ordinários e de prática científica. Por vezes, tais alegações descrevem o grau de suporte probabilístico que um conjunto de evidências oferece a uma determinada hipótese ou, não raramente, a frequência relativa de um tipo de evento ocorrendo dentro de uma classe de referência ou, ainda, dado que certas condições físicas são o caso no mundo, a propensão de certas causas produzirem um tipo de evento em particular. Precisamos, assim, distinguir os variados conceitos de probabilidade que são empregados em tais contextos, uma vez que eles não refletem uma única interpretação ou um mesmo tipo geral de probabilidade.

Algumas observações iniciais são decisivas. Primeira, sob certo aspecto, a tipologia que Swinburne propõe tem um caráter *sui generis*, especialmente em relação às suas definições de probabilidade e suas divisões e subdivisões entre tipos distintos. Segunda, o aparato formal de probabilidade apresentado na última seção funciona, precipuamente, para tipos de probabilidade indutiva. Mais rigorosamente, embora esse seja o *calculus* padrão, Swinburne propõe uma axiomatização alternativa para o tipo de probabilidade mais importante de sua teoria: probabilidades lógicas. A sua axiomatização define probabilidades lógicas em termos de probabilidades condicionais primitivas. Em todo caso, é matéria de controvérsia se o *calculus* básico se aplica igualmente para probabilidades estatística, frequência relativa, e física ou natural, propensões. Não se pretende disputar isso no momento.[37] Terceira, há uma extensa literatura sobre interpretações e análises dos conceitos de probabilidade.[38] P. S. Laplace (1951 [1820]) é considerado um dos principais defensores de uma interpretação clássica, onde probabilidades podem ser calculadas somente se há um número simetricamente distribuído de possibilidades em um espaço de amostra. Para qualquer evento A, sob as suposições de que n = número total de possibilidades e r = número de ocorrências nas quais um evento A é o caso, $Pr(A) = \frac{r}{n}$. Quanto à interpretação subjetiva da função de probabilidades, Ramsey (1950 [1926]) e De Finetti (1964 [1937]) são os principais expoentes. Em tal modelo, probabilidades correspondem a graus subjetivos de crença de agentes e, para que tais agentes sejam racionais em suas crenças, tais graus precisam satisfazer restrições de racionalidade

[37] Mais informações em Alan Hájek (2011, secs. 3.4 e 3.5).

[38] Seguem alguns textos informativos com uma excelente visão geral e crítica sobre o assunto: Donald Gillies (2000), D. H. Mellor (2005) e Alan Hájek (2011).

e coerência probabilística. Para tal, a defesa do argumento do *Dutch Book* tem sido a estratégia tradicional.[39] A propósito, muitos Bayesianos contemporâneos e epistemólogos formais têm se inspirado na proposta de explicar graus de crença pelo recurso do cálculo probabilístico. Veremos, contudo, que probabilidades subjetivas têm um sentido muito específico na teoria de Swinburne. Em relação à interpretação lógica da função de probabilidade, as teorias de Carnap[40] (1962) e Keynes (1921) estão entre as mais destacadas. O grau de confirmação de h condicional em e depende do grau de probabilidade lógica que e fornece a h. Proponentes dessa interpretação usualmente alegam que há um modo puramente objetivo de determinar a probabilidade de proposições. Bayesianos objetivos — como o próprio Swinburne —, que pretendem determinar probabilidades intrínsecas de proposições por critérios *a priori*, podem se sentir seduzidos pelas propostas de Carnap e Keynes. Por fim, vale ressaltar que usualmente interpretações propensitivas e frequentistas, a primeira defendida por Popper (1959) e a segunda por Reichenbach (1949) e von Mises (1957), têm sido classificadas como teorias concorrentes na explicação de um mesmo tipo de probabilidade, nomeadamente, probabilidades físicas[41] (*chance*). Surpreendentemente, Swinburne entende que as duas interpretações, na verdade, correspondem a probabilidades de natureza distinta e reserva um lugar especial a cada uma delas.

3.2.1 Probabilidade Física ou Natural

Para Swinburne (2001, p. 56-57; 2002, p. 2-3; 2013, p. 230), probabilidades físicas (ou naturais) são explicadas em termos de propensões. Se um determinado

[39]Como veremos no próximo capítulo, o argumento do *Dutch Book* é a posição *default* entre Bayesianos. Existem, entretanto, outros argumentos a favor da tese de que agentes devem obedecer ao cálculo de probabilidades, conhecida como a norma de coerência probabilística ou probabilismo. Os principais são o argumento baseado no teorema representacional (Patrick Maher, 1993 e 1997), o argumento da acurácia dos graus de crença (James Joyce, 1998 e 2009) e, mais recentemente, a teoria de utilidade epistêmica (Richard Pettigrew, 2013 e 2016 (*forthcoming*)), esta última inspirada na proposta da acurácia dos graus de crença de Joyce.

[40]Por sinal, R. Carnap (1962, cap. 2, p. 19) distinguiu somente dois conceitos de probabilidade: $tipo_1$ e $tipo_2$. O primeiro, probabilidade lógica, é o grau de confirmação de uma hipótese condicional em uma evidência, ou corpo evidencial, e o segundo, *chance*, é a frequência na qual um tipo de evento ocorre a longo prazo ou em uma longa sequência.

[41]Um estudo contemporâneo abrangendo outras teorias sobre *chance*, ou probabilidades físicas, foi desenvolvido por Toby Handfield (2012).

conjunto de condições físicas ou naturais[42] $C_1, C_2, \cdots, C_n$ torna fisicamente necessária a ocorrência de um evento D de um certo tipo em t, então a probabilidade física de D em t é 1. Por conseguinte, assumindo que $Ch(\,\cdot\,)$ é uma função de probabilidade física[43] (*chance*), $Ch_t(D \mid C_1, C_2, \cdots, C_n) = 1$. Se $C_1, C_2, \cdots, C_n$ tornam impossível fisicamente a ocorrência de um evento M de um certo tipo em t, então a probabilidade física de M em t é 0. Dessa maneira, $Ch_t(M \mid C_1, C_2, \cdots, C_n) = 0$. Agora, se $C_1, C_2, \cdots, C_n$ têm uma propensão em produzir um evento N de um tipo determinado em t — mas não o tornam nem necessário, nem impossível —, então a probabilidade física de N em t é ε tal que $0 < \varepsilon < 1$. Neste último caso, o grau de probabilidade física do evento N pode ser qualquer valor de uma diversidade de valores intermediários do intervalo $[0, 1]$, mas diferente de 0 e 1. Assim sendo, $Ch_t(N \mid C_1, C_2, \cdots, C_n) = \varepsilon$.

Tipicamente, probabilidades físicas são consideradas como propriedades ou características do mundo independentes objetivamente do fato de um agente crer ou saber a seu respeito.[44] Por exemplo, a probabilidade física de um isótopo de rádio decair dentro de um intervalo de 1.600 anos é .5, que corresponde à meia-vida dos seus átomos no processo de decaimento radioativo. Essa probabilidade descreve uma propensão natural do mundo em produzir esse evento e podemos ou não saber disso. Outro ponto central diz respeito ao fato de que tais propensões físicas ou naturais são entendidas por Swinburne (2002, p. 3) como propensões nas quais certas condições e causas operam na produção de um evento. Retomando o exemplo acima, Swinburne (2001, p. 57) sugere que no primeiro caso D é predeterminado a ocorrer em t, no segundo M é predeterminado a não ocorrer em t e, no último cenário, N tem uma tendência ou disposição física intermediária a ocorrer no mundo em t.

Uma observação crítica merece nossa atenção. De acordo com a distinção

[42]Estamos usando o termo 'condições' em um sentido bastante lato. Ele pode incluir tanto um conjunto de eventos e leis naturais, ou estados de coisas do mundo, ou causas anteriores que produzem um certo tipo de evento. Mas não é importante avaliarmos isso agora.

[43]Assumimos anteriormente que proposições são os objetos principais das funções de probabilidade, qualquer que seja o tipo. Contudo, uma vez que normalmente atribui-se probabilidades físicas a eventos, faremos uma concessão aqui, ou seja, vamos encapsular eventos como objetos da função de probabilidades físicas. Poderíamos, da mesma maneira, assumir que uma proposição h descreve a ocorrência de um evento D em t e, omitindo o condicionante da função, representar isso simplesmente como $Ch_t(h)$.

[44]Ver D. H. Mellor (2005, cap. 1) e Toby Handfield (2012, cap. 1).

feita por D. Gillies (2000, cap. 6, p. 126), teorias propensitivas podem ser teorias que descrevem propensões em que um tipo de evento ocorre em uma longa sequência (*long-run*) ou teorias nas quais propensões são relativas à ocorrência de um evento de um único caso (*single-case*) em uma certa situação. As primeiras supõem condições que se repetem em uma longa sequência de ocorrências de eventos de um tipo específico. Para tais teorias, propensões geram frequências relativas que têm valores próximos de probabilidades. Para as últimas, propensões não geram necessariamente uma longa sequência, tampouco produzem frequências relativas. Suponha, mais uma vez, um mecanismo justo de jogadas de uma moeda. Por um lado, teorias propensitivas do primeiro modelo consideram que há uma propensão do mecanismo produzir o resultado de *coroa* com uma frequência relativa próxima de $\frac{1}{2}$ em longa sequência de jogadas. Por outro lado, teorias propensitivas do segundo modelo enunciam unicamente que a propensão do mecanismo produzir o resultado de *coroa* tem o valor de probabilidade de $\frac{1}{2}$ em um certo tempo. Swinburne não se envolve em tal debate a respeito de teorias de propensão de forma explícita. Entretanto, as suas alegações em passagens diferentes podem nos causar alguma confusão sobre qual é a sua preferência entre os dois modelos. Em *Epistemic Justification* (2001, p. 57), quando reporta-se aos *relata* de probabilidade físicas e propensões, Swinburne os descreve como 'um resultado (ou evento) em particular'.[45] Diferentemente, em *Mind, Brain, and Free Will* (2013, p. 230), ele os descreve como 'um evento de um certo tipo'.[46] Uma melhor interpretação de sua teoria, todavia, revela que (1) Swinburne claramente distingue propensões de frequências relativas e, baseando-se nelas, oferece explicações divergentes para tipos distintos de probabilidade, e (2), para ele, frequências relativas correspondem a probabilidades estatísticas, não probabilidades físicas. Mais precisamente, Swinburne (2001, p. 58) reivindica que propensões, mensuradas por probabilidades físicas, têm 0 ou 1 como valores (ou algo muito próximo disso). Isso porque as condições físicas iminentes ao momento no qual o evento é ge-

[45]'First there is physical (or natural) probability [...] This is a measure of the extent to which **some particular outcome** is predetermined by its causes at some earlier time' (2001, p. 57).

[46]'If there is a natural probability of 1 at a time t that **an event of a certain kind** will occur, then the state of the world at t naturally necessitates its occurrence' (2013, p. 230).

rado são determinísticas.[47] No exemplo acima, portanto, a probabilidade de $\frac{1}{2}$ seria melhor interpretada como uma *ratio* na qual um certo tipo de evento é produzido pelo mecanismo de jogadas em uma longa sequência repetível. Voltaremos nossa atenção, nesta altura, às probabilidades estatísticas.

3.2.2 Probabilidade Estatística

Em linhas gerais, probabilidades estatísticas correspondem a uma *ratio*, ou a uma proporção, de um tipo de evento produzido por um processo repetível em uma longa sequência. É a frequência relativa na qual um tipo de evento ocorre em uma classe de referência. Ou, de modo similar, é a frequência de um atributo com uma certa propriedade relativamente a uma classe de referência. Assim, outras coisas sendo iguais, a probabilidade estatística de um dispositivo de jogadas de um *dado* gerar o resultado de qualquer número ímpar é ou se aproxima de $\frac{1}{2}$.

No entanto, frequências relativas não são determinadas exclusivamente por um caso único (*single-case*). A frequência de um resultado acontecer é maior ou menor de acordo com as suas ocorrências em uma classe de referência, a saber, depende do número de instâncias que contam como favoráveis de um total de resultados. Em uma versão simples de frequentismo finito,[48] se n é o número de instâncias favoráveis a um certo tipo de evento A, as ocorrências de A em uma longa sequência, e m é número total de ocorrências correspondentes a uma classe de referência, então a probabilidade estatística de A é simplesmente a proporção de $\frac{n}{m}$, ou tende a ser isso. Ou, pelo recurso de probabilidade condicional, podemos representar frequências relativas da seguinte forma. Considere o mecanismo de jogadas de uma moeda e uma classe constituída somente de dois resultados, ou seja, $\Omega = \{cara, coroa\}$. Agora, condicional no fato de que se trata de um resultado de tal classe de referência Ω, procuramos pela probabilidade estatística de ocorrências de *coroa*, *i.e.*

[47] Em outra passagem, Swinburne (2002, p. 3) atribui graus probabilísticos intermediários de propensão a eventos microfísicos em um dado tempo t — tipicamente aqueles que são descritos em mecânica quântica, pois são situações de indeterminação física — e graus de propensão 1 ou 0 (ou muito próximos de tais limites superior e inferior) a eventos macrofísicos em t.

[48] Alan Hájek (2011, sec. 3.4) reconhece semelhanças entre a interpretação clássica de probabilidades e o frequentismo finito, mas alerta que a primeira considera todos resultados possíveis, ao passo que a segunda inclui somente ocorrências atuais.

$Pr(\{coroa\} \mid \{cara, coroa\})$. Portanto, se são 500 instâncias de *coroa* dentro de 1.000 ocorrências no total, temos a *ratio* de $\frac{1}{2}$, que corresponde a probabilidade estatística do resultado de *coroa* dentro da classe referida.

Acontece que, conforme sugestão de Swinburne (2001, p. 57-58), classes podem ser (1) atuais ou hipotéticas e (2) finitas ou infinitas. Usando o exemplo oferecido por ele, a proporção de eleitores de *New Hampshire* que votaram no candidato republicano nas eleições presidenciais do ano de 2000 nos Estados Unidos corresponde aproximadamente a 56%. Trata-se de uma classe atual, porque reflete um estado de coisas que de fato ocorreu, e finita, porque o número de membros da classe de referência é finito. Diferentemente, a proporção de números primos no conjunto dos números naturais $\mathbb{N}$ abrange uma proporção em uma classe infinita. Além disso, em uma classe hipotética, se um mecanismo fosse ativado para jogar um determinado *dado*, ele geraria uma proporção de resultados de números ímpares, em contraste com os resultados de números pares, em uma sequência total de ocorrências nessa classe; uma *ratio* de $\frac{1}{2}$ seria respectiva ao resultado produzido pelo mecanismo se tal e tal fosse o caso.

Alguns comentários e dificuldades devem ser observadas. Em primeiro lugar, proporções e *ratios* não parecem constituir um sentido forte e genuíno de probabilidade. Como observa Alan Hájek (2011, sec. 3.4), especialmente em relação às classes finitas de referência, proporções estão mais próximas de meras contagens do que de probabilidades. Vamos considerar a seguinte classe: $\Omega = \{$David Bowie, Barack Obama, Dilma Rousseff, John Lennon$\}$. A proporção de músicos em Ω é, efetivamente, $\frac{1}{2}$. Mas dificilmente diríamos que a probabilidade de um membro de tal classe ser músico é $\frac{1}{2}$. Em segundo lugar, frequências relativas hipotéticas em classes de referência infinitas têm sido uma alternativa em propostas frequentistas. Nessa perspectiva, probabilidades estatísticas são entendidas como frequências de amplitude infinita das quais nós podemos conceber hipoteticamente uma *ratio* que converge para um valor. São *limiting relative frequencies*.[49] Com efeito, em *Epistemic Justification*, Swinburne (2001, p. 58) assume que normalmente probabilidades

[49] Mais informações sobre propostas de frequentismo hipotético, problemas sobre classes de referência infinita e possíveis soluções em D. H. Mellor (2005, cap. 3) e Alan Hájek (2011, sec. 3.4).

estatísticas atuais são medidas por uma *ratio* em uma classe de referência finita e, em contrapartida, probabilidades estatísticas hipotéticas são medidas por uma *ratio* em uma classe infinita. Não obstante o caráter atrativo do frequentismo hipotético, uma séria desvantagem o acomete: uma mesma classe de referência infinita pode ter valores diferentes se a sua ordem é alterada. Suponhamos o conjunto (infinito) dos número naturais $\mathbb{N}^*$, isto é, $\mathbb{N}^* = \{x \in \mathbb{N} \mid x \neq 0\}$. Na ordem regular, a probabilidade estatística de obtermos um número par de $\mathbb{N}^*$ é, presumivelmente, $\frac{1}{2}$. Quanto mais números obtemos de $\mathbb{N}^*$, mais o valor converge para $\frac{1}{2}$. Se pegamos os dez primeiros números, conseguimos $\frac{1}{2}, \frac{1}{3}, \frac{2}{4}, \frac{2}{5}, \frac{3}{6}, \frac{3}{7}, \frac{4}{8}, \frac{4}{9}, \frac{5}{10}$; e assim sucessivamente quanto mais números obtivermos. Se mudamos a ordem para 3, 5, 7, 9, 2, 11, 13, 1, 15, 4, $\cdots$, o valor converge para $\frac{1}{5}$. A questão, então, é determinar algum critério que permita distinguir a ordem correta das incorretas, supondo que alguma delas detém alguma prerrogativa. Vale dizer que, advertidamente, Swinburne (2001, p. 58-59) discute esse problema e reconhece a sua força. Por ora, uma vez que não é o nosso principal objetivo oferecer uma resposta a ele, deixaremos a questão em aberto.

3.2.3 Probabilidade Indutiva

Quando uma proposição e torna uma proposição h mais provável do que $\neg h$, dizemos que e oferece alto grau de suporte probabilístico a h, tipicamente um grau ε tal que $.5 < \varepsilon \leq 1$. Normalmente, e é uma peça de evidência, ou um conjunto evidencial mais amplo formado por uma conjunção $e_1 \wedge e_2 \wedge \cdots \wedge e_n$, que contribui para a probabilidade de uma hipótese h. Neste caso, não estamos lidando com probabilidades estatísticas, nem com probabilidades físicas, mas com um tipo diferente de probabilidades: indutivas ou evidenciais.

Se e suporta h com grau máximo, então a probabilidade de h condicional em e é 1. Talvez uma instância última disso seja quando e acarreta h. Se $e \vDash h$, então $e \equiv (e \wedge h)$. Ora, se $e \equiv (e \wedge h)$, então $Pr(e) = Pr(e \wedge h)$ (teorema 3 do apêndice). Pela definição de probabilidade condicional, $Pr(h \mid e) = \frac{Pr(h \wedge e)}{Pr(e)}$. Uma vez que $(h \wedge e) \equiv (e \wedge h)$, substituindo $Pr(e \wedge h)$ por $Pr(e)$ na última fórmula, $Pr(h \mid e) = \frac{Pr(e)}{Pr(e)} = 1$, assumindo que $Pr(e) \neq 0$. Portanto, se $e \vDash h$, então

$Pr(h \mid e) = 1$. Se, ao invés disso, e suporta $\neg h$ com grau máximo, então a probabilidade de h condicional em e é 0. Pelo mesmo raciocínio, se $e \vDash \neg h$, $Pr(\neg h \mid e) = 1$. Assim, se $Pr(\neg h \mid e) = 1$, então $Pr(h \mid e) = 0$, pois $Pr(\neg h \mid e) + Pr(h \mid e) = 1$. Alternativamente, se e suporta h com algum grau intermediário, então $Pr(h \mid e) = \chi$ e $1 > \chi > 0$. Como *desideratum* epistêmico, χ deve ser um grau que se situa no subintervalo $(.5, 1]$, ou seja, $\{\chi \in \mathbb{R} \mid .5 < \chi \leq 1\}$. Quanto mais a probabilidade indutiva de h condicional em e se aproxima de 1, mais suporte e fornece a h, o que significa otimizar a probabilidade indutiva de h sobre e. Embora nem sempre tenhamos valores precisos, podemos comparar probabilidades indutivas de duas hipóteses mutuamente exclusivas h e $\neg h$ condicionadas em uma mesma evidência e. Se $Pr(h \mid e) > .5$, então $Pr(\neg h \mid e) < .5$. Portanto, e suporta mais h do que $\neg h$, isto é, $Pr(h \mid e) > Pr(\neg h \mid e)$. Se $Pr(h \mid e) < .5$ e $Pr(\neg h \mid e) > .5$, então e suporta mais $\neg h$ do que h. Logo, $Pr(h \mid e) < Pr(\neg h \mid e)$.

Podemos sustentar, por consequência, que probabilidades indutivas estão condicionadas em evidências. É uma propriedade fundamental e inerente à sua natureza. Mais especificamente, a probabilidade indutiva de h não está condicionada somente em uma evidência e ou em um determinado corpo evidencial constituído por $e_1, e_2, \cdots, e_n$. As evidências de fundo disponíveis, ou o conhecimento de fundo,[50] também podem desempenhar uma função no suporte de h. Como veremos em mais detalhes no último capítulo, o grau de confirmação probabilístico de uma hipótese h condicional em e é formalmente representado pela função $Pr(h \mid e \wedge k)$, onde k expressa tal evidência ou conhecimento de fundo. De acordo com a definição de Swinburne (2001, p. 104; 2004, p. 16), k diz respeito a tudo que assumimos como evidência sobre como o mundo funciona quando avaliamos uma hipótese h e antes de uma nova peça de evidência e ser descoberta ou considerada. Por exemplo, suponha que h é a hipótese de que um mordomo matou a esposa de um milionário e e_t representa um conjunto de evidências que um detetive obteve no local do crime. Assim, e_t é uma conjunção $((e_1 \wedge e_2) \wedge e_3)$: a evidência de uma testemunha que viu o mordomo próximo ao local na hora do crime, as evidências que o detetive encontrou,

[50] Tradução dos termos '*Background evidence*' e '*background knowledge*'. Por vezes, o termo '*background information*' é igualmente usado na literatura sobre confirmação Bayesiana.

a saber, as digitais do mordomo na arma e, na cena do crime, um fio de cabelo que combina com o seu DNA. Porém, não somente e_t seria considerada no suporte de h. Alegadamente, o detetive tem à sua disposição evidências sobre a alta confiabilidade de testes de DNA com amostras de fios de cabelo, de que um conjunto de digitais corresponde a uma única pessoa, e assim por diante. Estas últimas compõem o conhecimento ou as evidências de fundo k.

Entretanto, probabilidades indutivas são separadas em três subdivisões na teoria de Swinburne: subjetiva, epistêmica e lógica. Cada uma delas é definida de acordo com a capacidade que o agente epistêmico tem em realizar inferências, reconhecer relações de acarretamento, extrair consequências lógicas de evidências e identificar critérios corretos de probabilidade lógica. Sob certo aspecto, a sua classificação reflete uma gradação das competências e da *performance* cognitiva de agentes, desde uma capacidade ilimitada até uma mais restrita, e se eles operam ou não com padrões de correção para probabilidades indutivas. Podemos, desse modo, organizar as definições dos seus subtipos de probabilidade indutiva com o seguinte esquema abaixo (Swinburne, 2001, p. 62-71):

⟨**Probabilidades Lógicas**⟩: dizem respeito aos graus corretos de probabilidade indutiva. São probabilidades determinadas pelos critérios corretos de probabilidade lógica que agentes logicamente oniscientes são capazes de alcançar;

⟨**Probabilidades Epistêmicas**⟩: são graus de suporte probabilístico que evidências fornecem a hipóteses, mas são determinadas por agentes com capacidade lógica limitada, a despeito de tais agentes aplicarem os critérios corretos de probabilidade lógica;

⟨**Probabilidades Subjetivas**⟩: são graus de suporte probabilístico que agentes atribuem a hipóteses baseando-se em evidências, mas de acordo com seus próprios critérios subjetivos de probabilidade e com capacidade lógica limitada.

Para Swinburne (2001, p. 69-70), os tipos de probabilidade indutiva não são somente relativos à força da evidência. Como dissemos, eles também são relativos às capacidades dos agentes. Agentes com onisciência lógica são capazes de reconhecer possibilidades lógicas relevantes, extrair habilmente consequências dedutivas delas e atribuir graus corretos de probabilidade indutiva a proposições. Para qualquer agente logicamente onisciente S^*, se S^* crê que $(p \wedge (p \rightarrow q))$, então S^* formará a crença de que q, uma vez que ele conhece a regra de *modus ponens* e sabe usá-la de modo competente. Mas se q é um absurdo e S^* é racional, então S^* não vai acreditar que q, mas deixará de crer que p ou $(p \rightarrow q)$.[51] Se S^* reconhece que $e \vDash h$, então S^* será capaz de atribuir probabilidade 1 a h condicional em e. Igualmente, S^* será capaz de atribuir probabilidade 1 a tautologias e 0 a contradições. Embora isso envolva um alto grau de idealização e perfeição lógica, é justamente o que Swinburne pretende defender com probabilidades lógicas.[52] São as atribuições corretas (pelo uso de padrões corretos) que agentes realizariam se eles fossem logicamente oniscientes. Por assim dizer, é o ideal que agentes ordinários perseguem. Por várias razões, agentes ordinários têm capacidades mais limitadas e falham em atribuir corretamente valores de probabilidade, em deduzir consequências deles e, igualmente, são falíveis nas suas práticas inferenciais dedutivas. De qualquer forma, ainda que restritamente, eles podem empregar critérios corretos de probabilidade lógica. No entanto, não é suficiente que um agente obedeça ao cálculo de probabilidades de tal modo que suas atribuições sejam coerentes do ponto de vista probabilístico. Estas são restrições necessárias mínimas. Eles devem realizar avaliações de probabilidade indutiva em conformidade com os padrões corretos.

Os padrões ou critérios corretos de probabilidade lógica defendidos por Swinburne (2001, p. 80-102) são distribuídos em dois grandes grupos: (*i*) critérios *a priori* e (*ii*) critérios *a posteriori*. Critérios *a priori* são simplicidade e escopo; critérios *a posteriori* são o encaixe da hipótese com a evidência de fundo e o poder explanatório da hipótese. O critério de simplicidade é o mais relevante de sua teoria: outras

[51] Agradecimento ao Rodrigo Borges por chamar atenção para esse ponto.

[52] 'But that measure of inductive support that would be reached by a logically omniscient being (that is, one who knows what are all the relevant logical possibilities and knows what they entail, and has correct inductive criteria) is what I shall call **logical probability**' (2001, p. 64).

coisas sendo iguais, hipóteses mais simples são mais prováveis. Apesar disso, se duas hipóteses h_1 e h_2 são igualmente simples, outros critérios podem desempenhar função mais determinante. Em sentido rigoroso, melhor uma hipótese cumpre com os requerimentos dos quatro critérios — a saber, mais bem-sucedida ela é no seu desempenho global —, mais logicamente provável ela é. Vale ressaltar, ademais, que esse conjunto de critérios constitui o cerne da versão de Bayesianismo que Swinburne tem em mente. Uma avaliação mais precisa e extensa de tais critérios será fornecida no próximo capítulo.

Por último, Swinburne (2001, p. 103 e 2002, p. 6) sugere uma versão alternativa de axiomas para probabilidades lógicas. Ela é similar às propostas oferecidas por Popper (1959) e Rényi (1970) porque estabelece probabilidades condicionais como primitivas. Diferentemente, nos axiomas de Kolmogorov, probabilidades incondicionais ou categóricas são primitivas e probabilidades condicionais são definidas por um quociente das primeiras. Assim, para uma função de probabilidades de dois lugares $Pr(\cdot \mid \cdot)$, uma álgebra $\mathbb{F}$ tal que $p, q, r \in \mathbb{F}$ e $Pr(\cdot \mid \cdot) : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{R}$, temos os seguintes axiomas do *calculus** (onde '$\Box$' representa necessidade lógica):

- (1*) $Pr(\cdot \mid p) = \varepsilon$ e $\varepsilon \in [0, 1]$;
- (2*) $Pr(p \mid p) = 1$;
- (3*) $Pr(p \vee q \mid r) = Pr(p \mid r) + Pr(q \mid r)$ se $\Box\neg((p \wedge q) \wedge r)$ e $\neg\Box\neg r$;
- (4) $Pr(r \mid p) = Pr(r \mid q)$ se $\Box(p \leftrightarrow q)$;
- (5) $Pr(q \mid p) = 1$ se $\Box(p \rightarrow q)$;
- (6) $Pr(p \wedge q \mid r) = Pr(p \mid q \wedge r) \times Pr(q \mid r)$.

Em tal modelo formal, considerando que a é uma verdade necessária em $\mathbb{F}$, $Pr(\,\cdot\,)$ é equivalente a $Pr(\cdot \mid a)$. Tal axiomatização apresenta diferenças significativas em relação à axiomatização de Kolmogorov. Uma vantagem é que probabilidades condicionais podem ser definidas mesmo quando o denominador é igual a 0. Os axiomas (1*) e (2*) são variações dos axiomas de não-negatividade e normalização: qualquer função de probabilidade condicional tem um valor que se situa no intervalo $[0, 1]$ e qualquer proposição condicional em si mesma tem valor de probabilidade *maximum* 1. O axioma (3*), por seu turno, é uma versão sofisticada do axioma de aditividade finita. O axioma (4) diz que se é necessário que p se e somente se q, então, para qualquer r, a probabilidade de r condicional em p tem o mesmo valor que a sua probabilidade condicional em q. Por sua vez, o axioma (5) determina que se p implica estritamente que q, então a probabilidade de q condicional em p é 1; e $\Box(p \to q)$ equivale a $\neg\Diamond(p \wedge \neg q)$. No *calculus* tradicional, deriva-se como um teorema que se $p \vDash q$, então $Pr(q \mid p) = 1$. Sobre o último axioma, é possível demonstrá-lo como um teorema do *calculus* com o uso da definição de probabilidade condicional.[53] Todavia, ao menos que o contrário seja afirmado, manteremos os axiomas de Kolmogorov e as definições do *calculus* padrão subentendido em nossas discussões.[54]

3.2.4 Relações entre Tipos Diferentes de Probabilidade

Considere um atributo (ou evento) A de um certo tipo específico em uma classe de referência apropriada. Se a probabilidade estatística de A ser B em tal classe é ε e, para qualquer x, x é A, então a probabilidade lógica de x ser B é ε. Em outras palavras, a probabilidade lógica de x ser B é ε condicional no fato de que x é A e de que a probabilidade estatística de A ser B em uma determinada classe de referência é ε. Contudo, Swinburne (2001, p. 79) alerta que evidências adicionais podem alterar esse valor de probabilidade lógica. Usando o seu exemplo, se a probabilidade estatística de A ser B é $\frac{2}{3}$ e x é A, mas um agente confiável e com o seu aparelho perceptual funcionado adequadamente observa que não é o caso que x é B em t, então a probabilidade lógica de x ser B em t é 0. Analogamente, se a probabilidade física de um conjunto de condições $C_1, C_2, \cdots, C_n$ produzir a ocorrência de um evento M é ζ, supondo que uma lei natural descreve corretamente tais propensões, e as condições $C_1, C_2, \cdots, C_n$ são o caso no mundo, então a probabilidade lógica de M ocorrer é igualmente ζ, supondo nenhuma outra informação ou evidência adicional contrária.

No primeiro caso, temos uma formulação que se assemelha ao princípio de probabilidade direta em Ian Hacking (1965). Observe, no entanto, que probabilida-

[53] Ver teorema 10 disponível no apêndice.

[54] Mais sobre probabilidades condicionais e versões alternativas de axiomatização em John Earman (1992, cap. 2), Alan Hájek (2003 e 2011) e Jonathan Weisberg (2011).

des lógicas, ao invés de probabilidades subjetivas, devem corresponder a frequências relativas, supondo, evidentemente, a tipologia de probabilidades das seções acima. Swinburne (2001, p. 79) não se pronuncia explicitamente sobre esse princípio, mas claramente o emprega. Assumindo que h descreve um tipo de evento em uma classe de referência adequada e que $Fr(\,\cdot\,)$ é uma função para frequências relativas:

$$Pr(h \mid Fr(h) = \varepsilon) = \varepsilon$$

O segundo caso, entretanto, remete a uma versão modificada do princípio principal[55] proposto originalmente por David Lewis (1980 e 1994). Em particular, na versão de Swinburne (2001, p. 78), probabilidades lógicas, não graus de crença probabilísticos (*credences*), devem corresponder a probabilidades físicas. Assumindo que h descreve a ocorrência de um evento sob determinadas condições e que $Ch(\,\cdot\,)$ é uma função para probabilidades físicas:[56]

$$Pr(h \mid Ch(h) = \zeta) = \zeta$$

Se h ou $\neg h$, mas não ambas, estivessem entre os termos *condicionantes* da função, tais atribuições de probabilidade lógica assumiriam outros valores. Considere a frequência relativa de um mecanismo repetível de jogadas gerar o resultado de *coroa*. Suponha, além disso, que a sua probabilidade estatística seja $\frac{1}{2}$ em uma longa sequência considerada apropriada. Observe, agora, a seguinte situação:

$$Pr(coroa \mid coroa \ \wedge \ Fr(coroa) = .5) = 1$$

Mesmo que a probabilidade estatística de .5 de um resultado particular de *coroa* esteja como evidência na parte *condicionante* da função, o seu valor de probabilidade lógica é 1 porque a sua ocorrência também está como evidência em tal função. Do mesmo modo, se a ocorrência de um particular resultado de *não-coroa*

[55] Michael Strevens (1999) oferece uma discussão fecunda sobre tentativas de justificação do princípio principal, o que é para ele um tipo de princípio de coordenação probabilística.

[56] Para maior clareza, omitimos as variáveis que indicam instâncias de tempo na função.

estivesse como *condicionante* na função, então, trivialmente, a atribuição de probabilidade lógica para o resultado de *coroa* seria 0. Como dissemos, o próprio Swinburne admite que evidências adicionais podem alterar o valor das funções iniciais de probabilidade lógica, o que mostra que elas são não-monotônicas.

3.3 Probabilidade, Internalismo e Externalismo

Como mencionamos, Swinburne (2001, p. 56) alega que teorias internalistas e externalistas de justificação empregam tipos distintos de probabilidade para explicar o quão bem *grounds* de crenças estão adequados. O que significa probabilizar uma crença e conferir *status* epistêmico positivo a ela depende da concepção de justificação epistêmica e do conceito de probabilidade que se está assumindo. Swinburne (2001, p. 71-73) entende que enquanto teorias internalistas recorrem a algum tipo de probabilidade indutiva, teorias externalistas podem usar probabilidades físicas ou probabilidades estatísticas.[57] Em termos gerais e técnicos, mais provável um *ground* G, ou um conjunto de *grounds* G_1, G_2, $\cdots$, G_n, suporta uma crença B de S em t, mais adequação é conferida a G, ou a G_1, G_2, $\cdots$, G_n, e mais justificada está a crença B de S em t, considerando que ela está baseada em G ou G_1, G_2, $\cdots$, G_n; tipicamente, quando esse grau χ de probabilidade se situa no intervalo $(.5, 1]$ ou $.5 < \chi \leq 1$. Em outras palavras, uma crença B de um agente S está melhor justificada somente se ela está baseada em *grounds* adequados, isto é, se estes a tornam provável.[58] Adequação e justificação são, para Swinburne (2001, p. 11), uma questão de graus.[59] Antes de esclarecermos e avaliarmos os usos dos conceitos de probabilidade no contexto da adequação dos *grounds*, precisamos explicar algumas noções e definições centrais desse debate.

Em primeiro lugar, *grounds* podem ser razões ou evidências que suportam

[57]Embora Swinburne não explore essa possibilidade, talvez teorias externalistas de justificação possam empregar o conceito de probabilidade lógica oferecido por ele, uma vez que os padrões epistêmicos de agentes logicamente oniscientes parecem ser inacessíveis por agentes ordinários.

[58]Mais a clásula adicional de que não há qualquer contra-evidência ou derrotador que torne a crença inadequada ou improvável.

[59]Swinburne (2001, p. 9) entende que justificação pode ser avaliada em uma dimensão sincrônica, as razões e evidências que tornam uma crença justificada em t, e uma dimensão diacrônica, justificação sincrônica mais investigação adequada através do tempo, de t para t' e assim por diante.

crenças. Quando *grounds* são bons o suficiente, eles conectam crenças mais diretamente, ou as aproximam, da verdade. Ou seja, se eles têm algum papel justificador, então eles são conducentes à verdade. Conforme William Alston (2005, p. 83), pode-se classificar *grounds* em dois grandes grupos: doxástico ou não-doxástico. Por um lado, se são *grounds* de natureza doxástica, então crenças são *grounds* de outras crenças. Temos, assim, uma relação de embasamento entre itens proposicionais, uma vez que é parte constitutiva ou uma propriedade da crença ter conteúdo proposicional. Por outro lado, se *grounds* são de natureza não-doxástica, então experiências e sensações, não somente crenças, podem desempenhar essa função de embasamento. Quando uma determinada crença tem um *ground* de caráter não-doxástico, temos uma relação entre um item proposicional e um não-proposicional. Considerando a nossa suposição de que itens proposicionais são os objetos da função de probabilidade, focaremos em *grounds* doxásticos — ou razões ou, ainda, evidências proposicionais — daqui em diante.

A despeito de diferenças significativas entre teorias internalistas e externalistas, epistemólogos entendem que conducência à verdade é uma propriedade fundamental do conceito de justificação epistêmica.[60] Mais precisamente, justificação deve obedecer a duas cláusulas mínimas: produzir maximamente crenças verdadeiras e evitar ou minimizar crenças falsas. De todo modo, as condições pelas quais uma crença se torna justificada é matéria de disputa acirrada entre internalistas e externalistas. Em linhas gerais, para que uma crença seja justificada, o internalismo exige como condição necessária que o agente tenha algum acesso a tais condições ou fatores justificadores de sua crença, *via* introspeção ou reflexão, ou que tais fatores sejam estados mentais do agente.[61] O externalismo, contrariamente, rejeita que qualquer uma de tais condições requeridas por internalistas, acessibilidade ou fatores justificadores *qua* estados mentais, seja uma condição necessária para justificação de crenças. Para uma crença ser justificada, certas condições externas objetivas devem

[60] A esse respeito, ver William Alston (1989) e Matthias Steup (2005).

[61] Seguindo sugestão de Steup (2005, sec. 2.3), podemos distinguir internalismo de acessibilidade (R. Chisholm, 1977), um no qual agentes reconhecem os fatores e condições de justificação pela reflexão, de internalismo mentalista (E. Conee e R. Feldman, 2001), um no qual os fatores justificadores são estados mentais do agente.

ser satisfeitas, como o fato de uma crença ser produzida por um processo confiável. É suficiente que o processo de formação de crenças que gera a crença-alvo seja de um tipo confiável.

Quanto ao externalismo, uma primeira estratégia seria explicar a adequação dos *grounds* por meio de probabilidades físicas (*propensity-type*). Assim, *grounds* são mais adequados somente se eles tornam fisicamente provável a ocorrência de uma crença verdadeira ou, alternativamente, somente se eles têm a propensão física de gerar mais crenças verdadeiras do que falsas. Como vimos, propensões explicam probabilidades físicas de eventos naturais: a probabilidade de certas condições $C_1, C_2, \cdots, C_n$ produzirem a ocorrência de um evento L. Não é exatamente claro como propensões funcionariam em relação a crenças: quais são as condições requeridas para que um conjunto de *grounds* produza mais crenças verdadeiras do que falsas, se propensões dão origem a um caso único de crença ou se crenças são produzidas por propensões que geram frequências relativas. A propósito, como bem salienta William Alston (2005, p. 109), propensões não parecem ser características intrínsecas dos *grounds* de crenças. Em última análise, Swinburne (2001, p. 71-72) considera que eventos de média ou grande escala têm graus de probabilidade física 1 ou 0 (ou aproximam-se disso). Se crenças se enquadram nessa categoria de eventos e justificação depende da adequação dos *grounds*, todas as nossas crenças verdadeiras atuais seriam justificadas em grau máximo e nossas crenças falsas atuais em grau mínimo, o que não parece ser o caso.

Uma proposta mais apropriada pode ser formulada nos seguintes termos. Uma crença tem *grounds* mais adequados à medida em que um processo de formação de crenças de um tipo confiável torna esta crença-alvo estatisticamente provável.[62] Nessa proposta, confiabilidade é medida por uma *ratio* ou uma frequência relativa. Se a *ratio* ou a frequência relativa de geração de crenças verdadeiras é maior do que .5, significa que esse processo gera mais crenças verdadeiras do que falsas. Quanto maior a *ratio* de crenças verdadeiras, mais confiável é o processo. No entanto, frequências relativas precisam pertencer a uma determinada classe de

[62]Em grande medida, essa é a posição defendida por Alvin Goldman (1979 e 1986).

referência, finita ou infinita. Como ilustramos, classes finitas de referência parecem ser meras contagens e não probabilidades, ao passo que classes infinitas padecem do problema da ordem de sequência gerada; valores podem variar de acordo com a ordem. Ademais, qualquer que seja a classe à qual o processo corresponda, finita ou infinita, o problema é determinar a que classe de referência especificamente os *grounds* e as crenças pertencem. Mais claramente, a crença de S de que *há um livro sobre a mesa à sua frente* pode pertencer a uma classe mais ampla, crenças perceptuais de S em geral, ou a uma classe um pouco mais restrita, crenças perceptuais de S no seu quarto, ou, ainda, a uma classe mais específica, crenças perceptuais de S no seu quarto com iluminação forte.[63]

Quanto ao internalismo, a adequação dos *grounds* depende de quão provável indutivamente uma proposição é suportada. Uma crença de que p de S tem *grounds* mais adequados, e é melhor justificada, quando tais *grounds* a tornam provável com grau ε tal que $.5 < \varepsilon \leq 1$. Vimos, entretanto, que probabilidades indutivas no sentido de Swinburne são respectivas às capacidades do agente e ao uso correto dos padrões de probabilidade lógica: desde um nível mais restrito a um nível ideal de atribuições de probabilidade. Em relação às probabilidades epistêmica e subjetiva, alguns problemas podem ser colocados. Para quaisquer p e r, suponha que um agente S com capacidades lógicas limitadas creia que p com base na sua crença de que r em t. Assim, a sua crença em r funciona como um *ground* doxástico para a sua crença de que p em t. Para que sua crença em p tenha *grounds* adequados e seja melhor justificada, r deve tornar p provável. Se representamos a relação entre elas com uma função de probabilidade condicional, espera-se que $.5 < Pr_t(p \mid r) \leq 1$. Mas, então, poderíamos nos perguntar se a crença de S em r tem ou não *grounds* que a tornam provável, a saber, se há ou não *grounds* adequados para a crença em r. William Alston (2005, p. 97) alerta que mesmo que a probabilidade de p condicional em r seja alta, isso não deixa S em uma posição epistêmica positiva para crer que p com

[63] Além disso, conforme Vincent Hendricks (2006, p. 44-45), a *ratio* correspondente ao sucesso estocástico de produção de crenças verdadeiras pode incluir mundos possíveis normais (compatíveis com as crenças gerais que um agente tem sobre o mundo atual) ou ser restrito ao mundo atual. Caso inclua outros mundos possíveis, a questão é se isso pode ter algum tipo de impacto relevante sobre classes de referência. Mais informações sobre o confiabilismo e mundos possíveis normais em Alvin Goldman (1986, cap. 5) e Vincent Hendricks (2006, cap. 3).

base em r se S não tem razões fortes para a sua crença em r.

Além disso, um problema mais geral pode ser proposto contra a sugestão de Swinburne. A sua alegação, como dissemos no início desta seção, é de que justificação epistêmica depende da adequação dos *grounds*. Por sua vez, adequação dos *grounds* significa algum grau de suporte probabilístico no intervalo de $(.5, 1]$. Assim, para quaisquer p e r, se $Pr(p \mid r) = \chi$ e $Pr(p \mid r) > Pr(\neg p \mid r)$, então p é mais provável do que $\neg p$, ambas proposições condicionadas em r. Mas, alegadamente, muitos valores de tal subintervalo não constituem um *threshold* apropriado para o tipo de justificação relevante para conhecimento. Por exemplo, $\chi = .51$ é um grau muito ínfimo para alegar que, dada outras condições, os *grounds* são adequados e r oferece forte justificação para a crença de que p, embora p seja provável condicional em r; em outras palavras, .51 é um valor muito baixo para que a crença de que p tenha justificação e consista em um caso genuíno de conhecimento, evitando a gettierização.[64] Em termos de otimização, χ deve ser um valor muito próximo ou igual a 1, mas não qualquer um acima de .5.

Uma objeção ainda mais contundente pode ser formulada sobre as relações entre probabilidades epistêmica ou subjetiva, adequação dos *grounds* e justificação. No cálculo probabilístico, se a é uma verdade lógica, então $Pr(a \mid \cdot) = 1$ e se b é uma contradição lógica, então $Pr(b \mid \cdot) = 0$. Qualquer proposição adicionada ao lado direito da função de probabilidade condicional não diminui a probabilidade de a, tampouco incrementa a probabilidade de b. Em outros termos, qualquer proposição é irrelevante para a probabilidade de tautologias e falsidades lógicas. Por exemplo, a proposição *5 + 7 = 12* condicional na proposição de que *todos os cisnes são brancos* tem grau máximo 1 de probabilidade. Agora, considere que p representa a proposição *5 + 7 = 12* e r representa *todos os cisnes são brancos*. Uma situação na qual S crê que p com base na sua crença em r não é exatamente desejável epistemicamente, nem constitui um sentido forte de adequação de *grounds* e justificação epistêmica. Em última instância, a crença de S em r não é razão adequada para a sua crença em p porque r não desempenha nenhum papel epistemicamente relevante para sua

[64] Agradecimento ao Rodrigo Borges por chamar atenção para esse problema.

crença em p, embora $Pr(p \mid r) = 1$. Além disso, o cálculo de probabilidades requer que tautologias em geral, não somente as conhecidas, tenham grau máximo. Alegadamente, não temos justificação e *grounds* adequados para muitas verdades lógicas, como o teorema da categoria de Baire ou o último teorema de Fermat. Podem nos faltar razões ou evidências para crer justificadamente nesses teoremas, elas podem não oferecer tão boa justificação, os *grounds* podem não ser adequados, não obstante tais teoremas terem grau 1 de probabilidade lógica. Agentes ordinários com capacidades lógicas limitadas podem falhar de diversas e inumeráveis maneiras. Portanto, probabilidade alta *per se* não é suficiente para adequação dos *grounds* e, consequentemente, para justificação, especialmente quando falamos de agentes ordinários com capacidades limitadas, ou seja, se empregamos os sentidos de probabilidades epistêmica e subjetiva de Swinburne.[65]

No entanto, a definição de probabilidade lógica de Swinburne tem uma consequência intrigante. Probabilidades lógicas dizem respeito aos graus corretos de probabilidade indutiva que agentes logicamente oniscientes são capazes de atingir. Uma vez que um agente logicamente onisciente S^* reconhece possibilidades lógicas relevantes e realiza inferências competentemente, a questão é se S^* tem as razões corretas para crer que o último teorema de Fermat é verdadeiro e, baseando-se nelas, se é o caso que a justificação da crença de S^* em tal teorema tem grau máximo 1; um grau que corresponde à probabilidade de que esse teorema é verdadeiro e satisfaz o axioma de normalização ou certeza do cálculo de probabilidades. Neste caso, probabilidades e justificação estão em conformidade, embora unicamente para agentes logicamente oniscientes. Com efeito, trata-se de um experimento mental a partir das condições e definições propostas por Swinburne. Em todo caso, o ônus de associar um tipo de probabilidade indutiva a um padrão dessa natureza é se comprometer com um ideal inatingível para agentes ordinários.

[65]Nessa esteira, Luis Rosa recentemente sugeriu uma objeção na qual probabilidades não correspondem a um sentido de justificação pessoal, isto é, justificação somente quando o agente tem razões disponíveis para crer que p é verdadeira. Verdades lógicas têm probabilidade 1 condicional em qualquer razão ou evidência e, por vários motivos, o grau de justificação que o agente tem para crer em uma verdade lógica pode não corresponder a esse grau máximo de probabilidade: porque o agente falha em ter razões para essa crença ou as razões não oferecem alto grau de justificação, entre outros. Apesar dos vários problemas relativos a graus de justificação como probabilidades, uma proposta interessante pode ser encontrada em Tomoji Shogenji (2012).

4 Bayesianismo

A despeito dos problemas envolvendo o uso dos diversos conceitos de probabilidade como medidas para adequação dos *grounds* e graus de justificação epistêmica, um emprego bem-sucedido de probabilidades indutivas e do *calculus* pode ser encontrado no contexto de confirmação de hipóteses explanatórias. Para tal, será decisivo introduzir uma peça fundamental do aparato probabilístico, a saber, o teorema de Bayes. Depois de explicarmos os seus aspectos técnicos e o uso específico que Swinburne faz desse teorema, focaremos na distinção entre Bayesianismo subjetivo e Bayesianismo objetivo. Swinburne pode ser classificado como um Bayesiano objetivo, visto que dois dos seus critérios de probabilidade lógica funcionam como princípios *a priori* para atribuições de probabilidade indutiva. Apesar das diferenças entre essas duas doutrinas, vale destacar que o Bayesianismo *lato sensu* não é uma teoria restrita tão-somente ao teorema de Bayes. Ele engloba um conjunto mais amplo de princípios, ferramentas formais e restrições de racionalidade probabilística. Ademais, é tarefa deste capítulo analisar o conjunto de critérios de probabilidade lógica que Swinburne advoga: poder explanatório, evidência de fundo, escopo e simplicidade. Em seguida, discutiremos o princípio de indiferença, dois conceitos de confirmação (incremental e absoluto) e, no final, exploraremos problemas e objeções à concepção de Swinburne e ao programa teórico Bayesiano como um todo.

4.1 Teorema de Bayes

Apresentamos e discutimos o aparato básico do *calculus* no último capítulo. Resta-nos, agora, adicionar uma última engrenagem a esse maquinário: o teorema de Bayes. O nosso propósito é usá-lo como uma ferramenta da teoria de confirmação. Antes, no entanto, mostraremos alguns dos seus aspectos formais e discutiremos as suas diferentes versões contemporâneas. Elas podem ser deduzidas em poucos passos inferenciais e de forma bem simples da definição de probabilidade condicional e do teorema de probabilidade total.[66] Mas, de fato, a ubiquidade do teorema de Bayes

[66] Ver teoremas 12 e 13 no apêndice.

e do Bayesianismo é realmente extraordinária na ciência e na filosofia atual.[67] É importante observar, contudo, que foi um clérigo e matemático inglês do século 18, Thomas Bayes (c. 1701-1781), quem primeiramente ofereceu uma demonstração da versão original de tal teorema.[68]

Para quaisquer h, e, k e assumindo que $Pr(e \mid k) > 0$:

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

Em um primeiro cenário, h representa uma hipótese explanatória, e é uma nova evidência observacional e k é o conhecimento ou evidência de fundo assumido quando e e h são avaliadas. Assim, $Pr(h \mid k)$, a probabilidade de h antes de e ser adicionada como evidência no lado direito da função, e $Pr(e \mid k)$ são as probabilidades iniciais ou anteriores (*priors*), $Pr(e \mid h \wedge k)$ é o *likelihood*, a probabilidade de e condicional em h e k, e $Pr(h \mid e \wedge k)$ é a probabilidade posterior, que mensura o impacto de e sobre h. Da mesma forma, $Pr(h \mid k)$ pode ser entendida como a plausibilidade inicial de h e, obviamente, $Pr(h \mid e \wedge k)$ representa o valor de probabilidade da hipótese-alvo que se pretende alcançar após os *inputs* do lado direito da equação serem computados. Além disso, Swinburne (2002, p. 10; 2004, p. 69) prefere chamar $Pr(e \mid h \wedge k)$ de poder preditivo,[69] mas, para ele, é irrelevante se h foi formulada em t e em t' descobriu-se e, ou se e foi descoberta em t e em t' formulou-se h ou mesmo se a formulação de h e a descoberta de e foram ambas no mesmo tempo t.

Em um segundo cenário, no entanto, Swinburne (2001, p. 105; 2004, p. 69)

[67] Seguem alguns exemplos de sucesso do Bayesianismo em diversas áreas. Em teoria da decisão, Richard Jeffrey (1983b) e Mark Kaplan (1996) proporcionam análises fecundas do aparato Bayesiano, discutindo problemas envolvendo a racionalidade de preferências e escolhas. Colin Howson e Peter Urbach (2006) desenvolvem uma teoria articulada, respondendo a problemas clássicos e comparando-a com análises frequentistas, com enfoque na estatística. Em epistemologia formal, Luc Bovens e Stephan Hartmann (2003) aplicam métodos Bayesianos sobre os conceitos de coerência, confirmação, testemunho e confiabilidade.

[68] A título de curiosidade histórica, foi Richard Price o responsável pela publicação póstuma da obra de Thomas Bayes no ano de 1763 na *Royal Society of London*. O artigo original de Bayes, *An Essay towards solving a Problem in the Doctrine of Chances* (2002 [1763], p. 117-149), está disponível no volume *Bayes's Theorem* editado por Richard Swinburne, além de outros textos com aplicações do Bayesianismo. A propósito, John Earman (1992, cap. 1, p. 7-31) fornece uma excelente avaliação do texto original de Bayes.

[69] Apesar do seu emprego técnico nesse contexto, o termo *likelihood* tem o mesmo sentido de *probability* no uso corrente do idioma inglês. Por isso, Swinburne prefere o termo *predictive power*.

entende que e pode representar toda evidência contingente logicamente relevante para h e, consequentemente, k torna-se uma variável que diz respeito estritamente à evidência ou conhecimento tautológico de fundo. Neste caso, $Pr(h \mid k)$, a probabilidade inicial de h, é definida como probabilidade intrínseca de h e os critérios *a priori* de escopo e simplicidade desempenham papel fundamental sobre essa função de probabilidade; igualmente, $Pr(e \mid k)$ é definida como probabilidade intrínseca de e. Mas se alguma evidência contingente sobre como o mundo funciona ou é categoricamente for adicionada a k, então o quão bem h se combina com a evidência contingente em k terá impacto sobre a probabilidade de $Pr(h \mid k)$. Falaremos em mais detalhes dos critérios de probabilidade lógica de Swinburne na seção 4.4.

O teorema de Bayes nos permite avaliar comparativamente as probabilidades de hipóteses concorrentes. Considere uma partição de duas hipóteses: $\Omega = \{h, h'\}$. Para tais avaliações, vamos supor que $Pr(h \mid k) \neq 0$, $Pr(h' \mid k) \neq 0$ e $Pr(e \mid k) \neq 0$. Para um mesmo conhecimento de fundo k e evidência e, se os seus *likelihoods* são iguais, $Pr(e \mid h \wedge k) = Pr(e \mid h' \wedge k)$, então a probabilidade posterior de uma das hipóteses será maior do que a da outra se e somente se a probabilidade inicial da primeira for maior do que a da segunda. Em termos técnicos, se $Pr(e \mid h' \wedge k) = Pr(e \mid h \wedge k)$, então $Pr(h' \mid e \wedge k) < Pr(h \mid e \wedge k)$ se e somente se $Pr(h' \mid k) < Pr(h \mid k)$; alternativamente, se $Pr(e \mid h' \wedge k) = Pr(e \mid h \wedge k)$, então $Pr(h' \mid e \wedge k) > Pr(h \mid e \wedge k)$ se e somente se $Pr(h' \mid k) > Pr(h \mid k)$. Mas, se $Pr(e \mid h \wedge k) = Pr(e \mid h' \wedge k)$ e $Pr(h \mid k) = Pr(h' \mid k)$, considerando que $Pr(e \mid k)$ se mantém constante, então $Pr(h \mid e \wedge k) = Pr(h' \mid e \wedge k)$. Analogamente, se $Pr(h' \mid k) = Pr(h \mid k)$, $Pr(h' \mid e \wedge k) < Pr(h \mid e \wedge k)$ se e somente se $Pr(e \mid h' \wedge k) < Pr(e \mid h \wedge k)$, e assim por diante. Por último, quando os valores dos *priors* e *likelihoods* são diferentes, admitindo que $Pr(e \mid k) = \chi$ tal que $\chi \neq 0$, $Pr(h \mid e \wedge k) > Pr(h' \mid e \wedge k)$ se e somente se $Pr(h \mid k) \times Pr(e \mid h \wedge k) > Pr(h' \mid k) \times Pr(e \mid h' \wedge k)$. Em última análise, tais avaliações comparativas são consequências lógicas do próprio teorema de Bayes.

A primeira versão do teorema de Bayes que apresentamos é uma forma simplificada de uma formulação mais completa do teorema. O denominador da equação contém a probabilidade de e condicional em k, ou seja, $Pr(e \mid k)$. Mas como vimos no capítulo anterior, o teorema de probabilidade total nos possibilita uma análise mais minuciosa das partes constitutivas de $Pr(e \mid k)$. Assim, para uma partição de duas hipóteses concorrentes h e $\neg h$, onde $Pr(h \vee \neg h \mid k) = 1$, e supondo que $Pr(e \mid k) \neq 0$:

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

Ainda, podemos generalizar o teorema de Bayes para uma partição de hipóteses concorrentes $\Omega' = \{h_1, h_2, h_3, \cdots, h_n\}$, onde $Pr(h_1 \vee h_2 \vee h_3 \vee \cdots \vee h_n \mid k) = 1$ e supondo que $Pr(e \mid k) \neq 0$. Desse modo, para qualquer h_i de Ω':

$$Pr(h_i \mid e \wedge k) = \frac{Pr(h_i \mid k) \times Pr(e \mid h_i \wedge k)}{\sum_j [Pr(h_j \mid k) \times Pr(e \mid h_j \wedge k)]}$$

O denominador, portanto, é constituído pela soma dos produtos das probabilidades iniciais com os *likelihoods* de cada hipótese da partição Ω'. Por consequência, considerando que h_1 é a nossa hipótese-alvo, menor o produto do *prior* com o *likelihood* de cada uma das outras hipóteses concorrentes, $[Pr(h_m \mid k) \times Pr(e \mid h_m \wedge k)]$, maior a probabilidade posterior de h_1, $Pr(h_1 \mid e \wedge k)$, sob a suposição de que $m \neq 1$.

4.2 Bayesianismo Subjetivo *vs.* Bayesianismo Objetivo

Bayesianos subjetivos alegam que graus de crença (*credences*) de agentes doxásticos são probabilidades subjetivas.[70] Mas para que tais graus subjetivos de crença sejam racionais, agentes precisam satisfazer os axiomas do cálculo da teoria de probabilidade (axiomas de Kolmogorov) e as suas consequências lógicas. Ou seja, no sentido empregado por Swinburne, Bayesianos subjetivos defendem que atribuições racionais de probabilidade indutiva obedecem ao cálculo de probabilidades. O argumento *default* a favor de tal primeira restrição — a saber, conformidade com

[70]Frank P. Ramsey em *Truth and Probability* (1950 [1926]) e Bruno de Finetti em *Foresight: its Logical Laws, its Subjective Sources* (1964 [1937]) são as principais influências de uma tradição de interpretação subjetiva da função de probabilidade e do que se entende atualmente por Bayesianismo subjetivo. Versões contemporâneas de Bayesianismo subjetivo podem ser encontradas em Paul Horwich (1982), Richard Jeffrey (1983b e 2004) e James Joyce (1998, 2004 e 2009).

o cálculo, o que podemos chamar de *coerência probabilística* — é conhecido como *Dutch Book* ou argumento do contrato de perda garantida. Uma segunda restrição sobre tais atribuições diz respeito à atualização dos graus de probabilidade subjetiva através do tempo, de t para t'. Nesse sentido, o princípio de condicionalização estrita e a regra de Jeffrey, ou princípio de condicionalização não-estrita, fornecem um esquema formal de como agentes devem atualizar suas probabilidades quando novas evidências são obtidas. Assim, o argumento do *Dutch Book* refere-se a uma dimensão *sincrônica* de racionalidade probabilística,[71] ao passo que princípios de condicionalização são relativos a uma dimensão *diacrônica* de racionalidade probabilística. Mais precisamente, probabilismo é a teoria segundo a qual (i) crenças se dão em graus, isto é, crenças são entendidas como um fenômeno gradual e (ii) que afirma que tais graus, ou atribuições de probabilidade subjetiva, são racionais apenas quando satisfazem o cálculo, *i.e.* a norma de racionalidade e coerência probabilística. Por seu turno, Bayesianismo é a teoria que oferece um esquema formal, teorema de Bayes + princípios de condicionalização, de como agentes devem atualizar os seus graus de probabilidade subjetiva quando ganham novas evidências. Portanto, a primeira se ocupa de uma dimensão *sincrônica* e a segunda de uma dimensão *diacrônica*.[72]

Podemos nos perguntar: por que razão agentes doxásticos devem obedecer ao maquinário de probabilidades? Em resposta a tal questão, Bayesianos frequentemente têm alegado que se um agente S não satisfaz o *calculus*, então S está vulnerável a um contrato de perda garantida (ou *Dutch Book*) e isso é suficiente para torná-lo irracional. Um primeiro aspecto importante do *Dutch Book* é a alegação de que probabilidades subjetivas e quocientes de aposta (*betting quotient*) estão conectados. Destarte, o grau de probabilidade subjetiva ε de um agente S em uma proposição r corresponde à disposição de S em apostar sobre a verdade de r. Se S tem grau .6 de que r, então significa que, para uma aposta que paga o montante de $R\$1$ como prêmio, S está disposto a pagar $R\$.6$ a favor de r. Podemos dizer que S tem um quociente de aposta $\frac{.6}{1}$ em r, ou seja, a *razão* é composta do valor de aposta (no

[71]Bayesianos subjetivos comumemente aceitam uma restrição adicional: o princípio principal de David Lewis. Como vimos, esse princípio enuncia que graus probabilísticos de crença devem corresponder a probabilidades físicas (*chance*).

[72]Ver André Neiva (2015). Mais informações em Richard Pettigrew (2013 e 2015).

numerador) pelo *stake* que está em jogo (no denominador).

Podemos formular o argumento do *Dutch Book* como se segue.[73] Graus de probabilidade subjetiva de um agente S devem corresponder aos quocientes de aposta de S. Se os quocientes de aposta de um agente S violam o cálculo probabilístico, então há um *Dutch Book* contra S (envolvendo tais quocientes).[74] Se há um *Dutch Book* contra S, então S está vulnerável a um contrato de perda garantida. Se S está vulnerável a um contrato de perda garantida, então S é irracional. Portanto, supondo a correspondência entre os quocientes de aposta e as probabilidades subjetivas de S, conclui-se por silogismo hipotético que se os quocientes de aposta de um agente S, e consequentemente os seus graus de probabilidade subjetiva, violam o cálculo probabilístico, então S é irracional.[75]

Dessa forma, sob a suposição de que $bq(\ \cdot\ )$ é uma função de quociente(s) de aposta e $\mathbb{F}$ é uma álgebra ou campo de proposições sobre um conjunto de possibilidades $\Omega \neq \emptyset$, tal que $bq : \mathbb{F} \longrightarrow \mathbb{R}$ ($p \in \mathbb{F}$ e $r \in \mathbb{F}$), temos os axiomas de probabilidade aplicados a tais quocientes:

- (1**) $bq(p) \geq 0$;
- (2**) Se p é uma tautologia, $bq(p) = 1$;
- (3**) Se p e r são incompatíveis, então $bq(p \vee r) = bq(p) + bq(r)$.

Antes de construirmos cenários de violação dos axiomas, precisamos fazer algumas suposições iniciais. Primeiro, apostar *a favor de p* equivale a apostar *contra* $\neg p$ e apostar *a favor de* $\neg p$ equivale a apostar *contra p*. Segundo, embora envolva uma certa quantidade de idealização, e para que o experimento mental com os quocientes de aposta funcione, geralmente supõe-se agentes que consideram recompensas monetárias como as suas únicas utilidades em contextos de aposta. É o que se entende,

[73]Ian Hacking (2001, p. 165), Franz Huber (2009, p. 5-6) e Darren Bradley (2015, cap. 3, p. 32) constroem versões similares ao argumento oferecido aqui.

[74]Esta segunda premissa é conhecida como Teorema do *Dutch Book*. Entretanto, se os quocientes de aposta (e as probabilidades subjetivas) de um agente S satisfazem o cálculo probabilístico, então não há um conjunto de apostas que torna S vulnerável a perda garantida. Isso é conhecido como *Converse Dutch Book Theorem*, demonstração em Kemeny (1955, p. 268-269).

[75]Esse é o argumento do *Dutch Book* na sua formulação mais canônica. No entanto, existem outros tipos de *Dutch Book* para justificar aditividade contável, princípios de condicionalização, entre outros. Para um exame desses outros argumentos, ver Susan Vineberg (2011).

de acordo com Christensen (2004, p. 117), por agentes simples (*simple agents*).

Se $bq(p) < 0$ de S, caso de violação do axioma de não-negatividade (1**), então S tem perda garantida. S entra em uma aposta com R\$$\varepsilon$ em p, mas paga a alguém R\$$\chi$ para levar seu bilhete. Significa que S abre mão de ganhar se p é o caso e tem perda monetária. E se $\neg p$ é o caso, S perde, pois apostou em p. No caso de violação do axioma de normalização (2**), temos duas possibilidades. Suponha que p é uma tautologia e $bq(p) > 1$. Significa que S paga R\$$\varepsilon$ por um bilhete de aposta com um prêmio de R\$$\chi$, mas $\varepsilon > \chi$. S tem perda garantida, pois o prêmio não cobre o valor apostado. Agora, no caso de $bq(p) < 1$ e $bq(p) = \varepsilon$, S também está disposto a apostar contra p (a favor de $\neg p$). Assim, o seu quociente em $\neg p$ é $bq(\neg p) = 1 - \varepsilon$. Um dono de casa de apostas poderia oferecer um *Dutch Book* a S com base em seu quociente contra p. Se o *stake* fosse R\$1, S perderia R\$$1 - \varepsilon$, uma vez que p é uma tautologia e certamente ocorrerá.

Casos de violação do axioma (3**), o de aditividade finita, são mais engenhosos e interessantes. Supondo um conjunto de duas proposições incompatíveis ou mutuamente exclusivas p e r e que $\zeta < \varepsilon + \chi$, temos os seguintes quocientes de aposta de S: $bq(p) = \varepsilon$, $bq(r) = \chi$ e $bq(p \vee r) = \zeta$. Com base em tais quocientes de aposta de S, um dono de casa de apostas concebe um contrato com um conjunto de apostas e oferece a S, com prêmio de R\$1: (a) aposta *a favor de* p em R\$0.3; (b) aposta *a favor de* r em R\$0.3; (c) aposta *contra* $p \vee r$ em R\$0.5 $(1 - \zeta)$. Ou seja, aposta *a favor de* $\neg(p \vee r)$ ou *a favor de* $\neg p \wedge \neg r$. Neste caso, S aposta valores diferentes para uma mesma coisa. Uma vez que p e r são incompatíveis, ele deveria ter quocientes em p, r e $p \vee r$ tal que $\zeta = \varepsilon + \chi$, a saber, $bq(p \vee r) = bq(p) + bq(r)$. Tal conjunto de apostas, por conseguinte, ocasiona perda garantida a S nas três possibilidades de resultado, como mostra a coluna do saldo na tabela abaixo:

	Aposta (a)	*Aposta (b)*	*Aposta (c)*	*Saldo*
$p \wedge \neg r$	+ R\$0.7	− R\$0.3	− R\$0.5	− R\$0.1
$\neg p \wedge r$	− R\$0.3	+ R\$0.7	− R\$0.5	− R\$0.1
$\neg p \wedge \neg r$	− R\$0.3	− R\$0.3	+ R\$0.5	− R\$0.1

O argumento do *Dutch Book* é dedutivamente válido e as violações de S resultam em perda garantida. Mas alguns problemas e objeções podem ser oferecidos a esse tipo de argumento. Primeiro, o *Dutch Book* não concerne à racionalidade epistêmica. Em última instância, o que está em discussão no caso de um agente não violar o maquinário formal de probabilidades diz respeito à racionalidade pragmática, a saber, o agente é racional do ponto de vista *pragmático* se não aceita apostas que o levam à perda garantida. O *Dutch Book* revela no máximo uma incoerência nos quocientes de aposta do agente.[76] Em grande medida, Swinburne (2001, p. 119-123) reconhece que o *Dutch Book* é um argumento sobre racionalidade pragmática. Segundo, há um argumento conhecido como *Czech Book*,[77] baseado em um suposto teorema do *Czech Book* e seu reverso, que basicamente diz o seguinte. Se os graus de probabilidade subjetiva de S, representados por quocientes de apostas, não satisfazem os axiomas do cálculo, então existe um conjunto de apostas que assegura ganho a S. Mas se os graus de probabilidade subjetiva de S obedecem aos axiomas do cálculo, então não existe um conjunto de apostas que assegura ganho a S. Desse modo, é um *desideratum* para o agente não obedecer ao cálculo: S não tem esse ganho se ele é probabilisticamente coerente. Terceiro, Swinburne (2002, p. 8-9) admite que probabilidades indutivas, quaisquer dos tipos da sua subdivisão, devem satisfazer o cálculo. Ainda que seja uma condição necessária, não é uma condição suficiente para atribuições corretas de probabilidade. Pode acontecer, todavia, que atribuições de probabilidade que obedecem ao cálculo de probabilidades sejam absurdas. Temos abundante evidência de que o amanhecer do Sol acontece todos os dias, isto é, em um intervalo aproximado de vinte e quatro horas. Se um agente S atribui probabilidade indutiva de .9 de que o Sol não nascerá amanhã e .1 de que ele nascerá, então S é coerente probabilisticamente, supondo que S não tem crenças de

[76] Um argumento mais promissor do que o *Dutch Book* é o argumento da acurácia dos graus de crença de James Joyce (1998 e 2009). O argumento de Joyce se ocupa estritamente com fins epistêmicos. O conceito de acurácia é medido por uma diferença entre duas funções, uma função de valor de verdade $w(\,\cdot\,)$ (*truth-value function*) e outra de graus de crença $Cr(\,\cdot\,)$, esta última não necessariamente probabilística. Tal diferença nos fornece um *score* $S(Cr, w)$ composto pelas duas funções e quanto maior o *score*, menos acurado é o grau de crença. Dessa maneira, uma função $Cr'(\,\cdot\,)$ domina uma função $Cr(\,\cdot\,)$ *se e somente se* $S(Cr, w) > S(Cr', w)$. Em linhas gerais, a estratégia de Joyce é demonstrar que funções probabilísticas de graus de crença não são dominadas por nenhuma outra função alternativa de graus de crença.

[77] Ver Alan Hájek (2009).

fundo sobre o comportamento correto do Sol. Mas embora essa atribuição satisfaça o cálculo, ela não é exatamente acurada. Não representa o estado de coisas do mundo adequadamente. Quarto, Swinburne (2005, p. 69-70) considera que as relações entre os termos nas funções de probabilidade são objetivas, não meramente relativas a um agente e aos seus próprios critérios de probabilidade subjetiva. Por assim dizer, ele esvazia do seu Bayesianismo esse forte componente subjetivista característico de versões mais radicais de Bayesianismo subjetivo.

Ainda nos falta falar dos dois princípios de condicionalização do Bayesianismo: (a) condicionalização estrita,[78] que trata especificamente de atualizações quando agentes tornam-se certos de uma determinada evidência, e (b) condicionalização não-estrita, proposta por Richard Jeffrey (1983b, p. 172), quando agentes revisam seu o grau de probabilidade em uma hipótese condicionalizando-o em uma evidência com valor não necessariamente maximal.

> **Condicionalização estrita**:[79] para quaisquer proposições **h** e **e**, se um agente **S** com probabilidade inicial $Pr_t(h \mid e) = \chi$, supondo que $1 > Pr_t(e) > 0$, recebe evidência de tal modo que torna-se certo de **e**, então **S** deve igualar o seu grau inicial em **h** condicional em **e**, $Pr_t(h \mid e)$, ao seu grau posterior em **h**, a saber, $Pr_t(h \mid e) = \chi = Pr_{t'}(h)$.

> **Condicionalização de Jeffrey**: para quaisquer proposições **h** e **e**, se um agente **S** com probabilidades iniciais $1 > Pr_t(h \mid e) > 0$ e $1 > Pr_t(e) > 0$, recebe evidência de tal modo que altera o seu grau inicial em **e** para $Pr_{t'}(e) = \chi$, então **S** deve revisar o seu grau inicial para um grau posterior em **h** tal que $Pr_{t'}(h) = [\chi \times Pr_t(h \mid e)] + [(1-\chi) \times Pr_t(h \mid \neg e)]$.

Podemos converter a equação da condicionalização de Jeffrey na seguinte fórmula: $Pr_{t'}(h) = [Pr_{t'}(e) \times Pr_t(h \mid e)] + [Pr_{t'}(\neg e) \times Pr_t(h \mid \neg e)]$, uma vez que $Pr_{t'}(\neg e) = 1 - Pr_{t'}(e)$. A condicionalização de Jeffrey tem a virtude teórica de explicar formalmente casos de atualização maximal $\chi = 1$ e não-maximal

[78] Ver William Talbott (2008, sec. 2) e Alan Hájek (2011, sec. 3.3).

[79] Omitimos a variável k das funções nos dois princípios pelo bem da simplicidade.

$0 < \chi < 1$. Dado que $\chi = 1$, temos $Pr_{t'}(h) = [1 \times Pr_t(h \mid e)] + [0 \times Pr_t(h \mid \neg e)]$ e, portanto, $Pr_{t'}(h) = Pr_t(h \mid e)$. A regra ou condicionalização de Jeffrey acima se aplica a uma partição de duas proposições: $W = \{e, \neg e\}$. Contudo, podemos generalizá-la para um conjunto mais amplo de proposições mutuamente exclusivas e conjuntamente exaustivas. Assim, para uma partição $W' = \{e_1, e_2, \cdots, e_n\}$, supondo que $Pr_t(e_i) \neq 0$ e $e_i \in W'$ ($i = 1$, $i = 2$, $\cdots$, $i = n$), segue-se que $Pr_{t'}(h) = \sum_{1 \leq i \leq n}[Pr_{t'}(e_i) \times Pr_t(h \mid e_i)]$.[80]

A princípio, Swinburne (2001, p. 246-248) aceita os dois princípios de condicionalização, mas com algumas restrições.[81] A primeira delas é que condicionalização deve ser realizada não somente em relação a uma nova peça de evidência contingente e_m, mas levando em consideração toda evidência disponível, a saber, todo conhecimento de fundo k + o conjunto de evidência já adquirido $\{e_1, e_2, \cdots, e_n\}$, dado que $n \neq m$. A segunda restrição diz que condicionalização pode ser feita com valores corretos de probabilidade. Mais precisamente, Swinburne (2001, p. 247) afirma: 'condicionalização é correta se e somente se eu estou operando com valores corretos de probabilidade'.[82] Por exemplo, usando o princípio de condicionalização estrita, se $Pr_t(h \mid e \wedge k) = .9$, $Pr_t(h \mid k) = \chi$, tal que $\chi < .9$, e nos tornamos certos de e, então, supondo que esses valores correspondem aos graus corretos de probabilidade indutiva, o novo grau de probabilidade em h deve ser .9, $Pr_{t'}(h \mid k) = .9$.

Geralmente, Bayesianos subjetivos não determinam restrições adicionais sobre probabilidades iniciais, os *priors*. Em situações que possuímos unicamente a informação de que uma partição é constituída de n possibilidades, qualquer atribuição de probabilidade inicial que satisfaça o cálculo, situando-se no intervalo de $[0, 1]$,

[80] Para qualquer e, se $Pr_{t'}(e) = 1$ após condicionalização, então esse valor mantém-se estável mesmo após sucessivas atualizações sobre evidências adicionais r, s, $\cdots$. Para qualquer r, usando condicionalização estrita e considerando que $Pr_{t'}(r) \neq 0$, $Pr_{t''}(e) = Pr_{t'}(e \mid r) = \frac{Pr_{t'}(e \wedge r)}{Pr_{t'}(r)} = 1$. O problema é que a probabilidade de e não pode ser diminuída de 1, ou seja, a sua certeza não pode ser perdida; e isso parece valer para ambos os princípios quando $Pr_{t'}(e) = 1$. Não discutiremos esse problema aqui, mas, em resposta a esse desafio, Bayesianos têm formulado princípios alternativos que acomodem o fato de que certeza pode ser perdida. Ver Timothy Williamson (2000, cap. 10) e Michael Titelbaum (2013) a esse respeito.

[81] Severas críticas aos princípios de condicionalização têm sido oferecidas por Bayesianos objetivos, que têm adotado, ao invés disso, o princípio de máxima entropia. A esse respeito, ver Jon Williamson (2010, cap. 4, p. 82-85).

[82] Na nota adicional J de *Epistemic Justification*: 'Conditionalizing is right if (and only if) I am operating with correct probability values' (2001, p. 247).

é racional. O problema para Bayesianos subjetivos é identificar sob que condições e em que momento os *priors* são determinados, um ponto de partida não arbitrário a partir do qual novas evidências são condicionalizadas. Por isso, Bayesianos objetivos têm procurado estratégias nas quais probabilidades iniciais podem ser restringidas. Apesar das diferentes versões,[83] podemos caracterizar o plano geral do Bayesianismo objetivo. Diferente de Bayesianos subjetivos, Bayesianos objetivos têm alegado que princípios e critérios *a priori* podem ser corretamente impostos sobre probabilidades iniciais, os *priors*, e, sobretudo, distribuições simétricas de probabilidade podem ser feitas sobre o espaço de possibilidades em tais situações de ignorância ou ausência de evidência sobre qual das n possibilidades é favorecida. Em especial, os critérios de escopo e simplicidade, assim como o princípio de indiferença, desempenham a função de princípios *a priori* na versão defendida por Swinburne. A ideia é que por meio desses princípios e critérios de natureza *a priori* estamos autorizados a determinar objetivamente probabilidades iniciais ou os *priors*.[84]

4.3 Explicação Científica e Explicação Pessoal

Antes de prosseguirmos com os critérios e princípios da sua teoria, parece adequado esclarecer o que Swinburne entende por hipóteses explanatórias.[85] Em sentido lato, uma hipótese explanatória pretende explicar as causas e os princípios de um conjunto de eventos ou fenômenos dos quais nós observamos. Tais hipóteses explanatórias podem envolver explicações de eventos observacionais do passado, de causas que operam subjacentemente a fenômenos dos quais observamos e, presumivelmente, permitem-nos tirar consequências sobre o que acontecerá no futuro.

[83]Para citar duas versões de Bayesianismo objetivo: a de E. T. Jaynes (2003) e, mais recentemente, a de Jon Williamson (2010).

[84]Mesmo que um valor númerico preciso χ não seja conferido a probabilidade inicial de h, William Talbott (2008, sec. 4.2) destaca que Bayesianos objetivos podem endossar uma estratégia segundo a qual o grau racional de probabilidade de h se situa em um intervalo $[a, b]$ tal que $[a, b] \subset [0, 1]$.

[85]No contexto de discussão voltado à avaliação dos argumentos a favor da existência de Deus de *The Existence of God* (2004), Swinburne discute minuciosamente diferenças entre os tipos de explicação, argumenta que explicação pessoal não é redutível à explicação científica, tipifica gradações de explicações, desde uma completa a uma absoluta, entre outros assuntos. Aqui, no entanto, vamos nos deter aos aspectos mais essenciais e básicos a respeito desse conceito, uma vez que a nossa atenção está destinada mais diretamente aos critérios de probabilidade lógica de Swinburne e à teoria Bayesiana de confirmação.

De acordo com a sugestão de Swinburne (2001, p. 74-75 e 2004, p. 23), existem dois padrões distintos de explicação: explicação científica e explicação pessoal. Quanto à explicação científica, podemos ainda distingui-la em dois subtipos: explicação plena e explicação parcial. Esses dois subtipos de explicação científica de Swinburne (2001, p. 75 e 2004, p. 26-27) estão baseados, respectivamente, nos conceitos de explicação nomológico-dedutiva e de explicação estatística-indutiva de Carl Hempel (1965). Se um evento E é dedutivamente acarretado por um conjunto de eventos ou estados de coisas, as condições iniciais $C_1, C_2, \cdots, C_n$, e por uma lei L, ou um conjunto de leis naturais $L_1, L_2, \cdots, L_n$, então trata-se de uma explicação plena. Em outras palavras, L e $C_1, C_2, \cdots, C_n$ explicam plenamente E. Assim, de um conjunto de condições iniciais $W = \{C_1, C_2, \cdots, C_n\}$, podemos ter um evento C_i, tal que $C_i \in W$, como a principal causa da ocorrência de um evento E. Normalmente, leis naturais descrevem generalizações como 'todos os As são Bs' e, em conjunção com as condições iniciais, são entendidas como a razão ou o motivo pelo qual o evento ocorre no mundo. Quando as condições iniciais $C_1, C_2, \cdots, C_n$ e uma lei natural L explicam parcialmente E, dizemos que $C_1, C_2, \cdots, C_n$ e L tornam E provável. Neste caso, L é uma lei probabilística como '$x\%$ de As são Bs'. Desse modo, a hipótese que descreve a ocorrência de E é tornada indutivamente provável pela evidência de que L e $C_1, C_2, \cdots, C_n$ são o caso no mundo e porque essa lei e tais condições estão operando em determinadas circunstâncias. De acordo com Swinburne (2004, p. 27), o objeto de explicações científicas não é restrito a eventos particulares. Ele também pode incluir leis mais específicas. Leis mais gerais podem explicar outras leis. Supondo certas condições $C_1, C_2, \cdots, C_n$, L_1 explica uma outra lei L_2 se L_2 é uma consequência dedutiva de L_1 ou é tornada provável por L_1.

No caso de explicações pessoais, um agente racional S é a causa da ocorrência de um evento E, ou seja, E é explicado em virtude de S ter desejado e exercido alguma ação intencional. Voltando à nossa discussão do capítulo 2, crenças de meios a fins, propósitos últimos e desejos de S funcionam como as razões pelas quais S, inclinado a escolher um curso de ação, realiza intencionalmente uma ação A_1. Se um agente S tem o propósito último de levantar a sua mão, motivado por algum desejo ou inclinação em agir de tal maneira, então provavelmente S levantará a sua mão. Pode ser o caso, porém, que S tenha um propósito último forte de atingir O_1 mediante uma ação A_1, mas, devido às circunstâncias, falhe em realizar A_1 e, por consequência, em alcançar O_1. Assim, embora S tente realizar A_1 e atingir O_1, S é impedido de alguma maneira. Por exemplo, ao tentar levantar a mão, alguma pessoa o impede. Em tais circunstâncias, Swinburne (2004, p. 35) diz que o evento é somente parcialmente explicado pela intenção de S. O agente tem o propósito último, tenta executá-lo por uma ação intencional, mas é obstruído. Como dissemos, pode haver uma cadeia causal de ações gerada por um propósito mais básico de S que visa um determinado fim. Ademais, Swinburne (2004, p. 37) destaca que uma explicação pessoal plena de E é fornecida quando um evento E é produzido por uma ação básica A_1 cujo propósito último é O_1 de S e, adicionalmente, quando S exerce os seus poderes básicos ou a sua causação por agência.[86]

Em suma, Swinburne (2004, p. 38-45) argumenta extensamente que explicação científica e explicação pessoal são de naturezas distintas e, por conseguinte, não se reduzem uma a outra. A primeira explica tipos de eventos que são causados por leis naturais e condições iniciais e a segunda, diferentemente, explica tipos de eventos produzidos por um agente racional que intencionalmente realiza uma ação. Ainda que isso seja de grande importância no contexto da sua argumentação teísta, essa caracterização geral nos é suficiente, uma vez que os nosso esforços atuais estão concentrados na sua teoria Bayesiana.

4.4 Critérios de Probabilidade Lógica

Falamos brevemente dos critérios de probabilidade lógica de Swinburne no capítulo anterior, particularmente quando definimos os subtipos de probabilidade indutiva. Neste momento, vamos definir mais acuradamente os seus quatro critérios e examinar cada um deles separadamente. Swinburne reserva um lugar de destaque ao

[86] Em *Mind, Brain, and Free Will* (2013), Swinburne argumenta a favor de uma forma de libertarianismo sobre metafísica da vontade livre, onde agentes causam intencionalmente suas próprias ações livres, e prefere uma teoria de causação e sobre leis naturais onde os *relata* causais não são eventos, mas sim substâncias, *i.e. substances, powers and liabilities account.* Mas isso não é importante agora.

seu critério de simplicidade. Ele é o mais importante critério de probabilidade lógica da sua versão de Bayesianismo. Em um conjunto de hipóteses mutuamente exclusivas e conjuntamente exaustivas tal que $\Omega = \{h_1, h_2, \cdots, h_n\}$, h_i de Ω é a mais provável, *ceteris paribus*, porque h_i é a mais simples.[87] Em adição ao critério de simplicidade, temos o critério de escopo. Em sentido estrito, ambos constituem os critérios *a priori* da teoria de Swinburne. Sobre os critérios *a posteriori*, temos o poder explanatório e o encaixe da hipótese com a evidência de fundo. Eles podem ser mais decisivos em situações nas quais duas ou mais hipóteses de mesmo escopo são igualmente simples, sobretudo o critério de poder explanatório. Ou seja, tais critérios podem funcionar como fatores de compensação. Em todo caso, se uma hipótese h_1 satisfaz melhor todos os quatro critérios do que h_2, h_3, $\cdots$, h_n, então a sua probabilidade lógica é maior do que as das suas competidoras. Em última análise, o ideal é que uma hipótese explanatória seja bem-sucedida na sua avaliação geral em relação aos critérios. Assim, em termos técnicos, de uma partição $\Omega = \{h_1, h_2, \cdots, h_n\}$, h_i é a mais provável ($h_i \in \Omega$) porque ela satisfaz melhor todos os critérios em conjunto do que qualquer outra hipótese de Ω.

4.4.1 Evidência de Fundo

Outras coisas sendo iguais, melhor uma hipótese h se encaixa com a evidência ou conhecimento de fundo k (*fitting with background evidence or knowledge*), mais provável é h. Ou, dito de outro modo, melhor o encaixe de h com k, maior a probabilidade de h condicional em k, a saber, maior a sua probabilidade inicial $Pr(h \mid k)$. Esse critério é considerado por Swinburne (2001, p. 81 e 2004, p. 53) como uma questão *a posteriori*, uma vez que ele depende do conhecimento ou evidência de fundo k. Depende da evidência total que é assumida sobre como o mundo funciona e da sua relação com uma hipótese explanatória. Se k é meramente tautológico, então a questão é sobre a probabilidade intrínseca de h. Se há alguma evidência contingente em k, então h precisa não somente se encaixar com o conhecimento tau-

[87] 'Other things being equal, a simpler hypothesis is more probably true and so the simplest hypothesis is the one most probably true' (Swinburne, 2001, p. 82).

tológico contido em k, não sendo inconsistente com ele, mas igualmente com essa parte evidencial contingente em k.

Considere que h represente a hipótese de que *João roubou um carro*. Suponha que a nossa evidência de fundo k contém a evidência de que *João foi preso várias vezes por roubos de carros*. Assim, h se encaixa melhor com k do que a hipótese contraditória $\neg h$. A sua probabilidade é maior condicional em k do que $\neg h$ condicional em k. Agora, se k contém a evidência de que *João nunca roubou um carro sequer*, então $\neg h$ se encaixa melhor com k do que h. Claro que poderíamos ganhar uma nova evidência e que derrota h (ou $\neg h$), mas a questão aqui é a respeito de quão bem h e k se combinam (ou $\neg h$ e k). De todo modo, se uma hipótese h é logicamente incompatível com a evidência de fundo k e com uma nova evidência e que se pretende explicar, então, presumivelmente, a probabilidade de h condicional na conjunção $k \wedge e$ é 0.

4.4.2 Poder Explanatório

O critério de poder explanatório é um segundo aspecto do Bayesianismo de Swinburne (2001, p. 80 e 2004, p. 69). Para quaisquer h e e, quanto maior o número de evidências contidas em e que h pretende explicar e maior o grau de probabilidade que h oferece como suporte a e, mais provável é o grau posterior de h. Significa que h explica mais evidências e realiza predições com maior grau de probabilidade sobre e. Assim, a sua probabilidade posterior, $Pr(h \mid e \wedge k)$, é maior. Em sentido estrito, o quociente $\frac{Pr(e \mid h \wedge k)}{Pr(e \mid k)}$, que é parte constituinte do teorema de Bayes, representa o que Swinburne define como poder explanatório de uma hipótese. Portanto, o poder explanatório é formado pela *razão* entre o *likelihood*, ou o poder preditivo, e a probabilidade de e condicional em k.

Swinburne é bastante breve sobre esse critério. Na verdade, ele é sucinto sobre todos os outros critérios, com exceção do critério de simplicidade. De qualquer forma, podemos tirar algumas consequências importantes. Supondo que $Pr(e \mid k) \neq 0$, se mantivermos constante a probabilidade inicial de h, $Pr(h \mid k)$, maior o valor de $\frac{Pr(e \mid h \wedge k)}{Pr(e \mid k)}$, maior a probabilidade posterior de h, $Pr(h \mid e \wedge k)$. Analogamente, quanto maior o *likelihood*, $Pr(e \mid h \wedge k)$, maior a probabilidade posterior de h, $Pr(h \mid e \wedge k)$, se $Pr(e \mid k)$ e $Pr(h \mid k)$ se mantêm constantes. Observe, além disso, que $Pr(e \mid h \wedge k)$ mede a probabilidade de e condicional em $h \wedge k$, isto é, o quanto h em conjunção com k prediz e. Se $(h \wedge k) \vDash e$, $(h \wedge k) \equiv ((h \wedge k) \wedge e)$. Dessa maneira, $Pr(h \wedge k) = Pr(h \wedge k \wedge e)$. Ora, $Pr(e \mid h \wedge k) = \frac{Pr(e \wedge h \wedge k)}{Pr(h \wedge k)}$. Assim, $Pr(e \mid h \wedge k) = \frac{Pr(h \wedge k)}{Pr(h \wedge k)}$. Portanto, se $(h \wedge k) \vDash e$, $Pr(e \mid h \wedge k) = 1$.[88] Neste caso, e é uma consequência de h e k, ambas predizem e com grau máximo. Ademais, como vimos no final da seção 4.1, quanto menor o produto da probabilidade inicial com o *likelihood* de cada uma das hipóteses concorrentes a h, maior a probabilidade posterior de h. Evidentemente, isso é o caso quando o produto da probabilidade inicial com o *likelihood* de h é maior do que os produtos desses termos relativos às outras hipóteses.

4.4.3 Escopo

Quanto ao conceito de escopo, Swinburne (2001, p. 82 e 2002, p. 12) o descreve como um critério *a priori*. Ele opera precipuamente sobre as probabilidades iniciais de hipóteses explanatórias. Normalmente, a relação entre o escopo e a probabilidade da hipótese é inversamente proporcional. Maior o escopo de uma hipótese, menor a sua probabilidade inicial. Menor o escopo, maior a sua probabilidade inicial. Uma hipótese é constituída de uma conjunção de alegações sobre a natureza de diversos tipos de entidades, eventos, propriedades do mundo e as suas relações. Maior é a extensão de objetos dos quais uma teoria se ocupa, maior é o risco de que ela afirme alguma falsidade sobre eles. Assim, grosso modo, o escopo da hipótese mede o quanto ela diz sobre o mundo. Por exemplo, uma hipótese h_1 tem maior escopo do que uma hipótese h_2 porque h_1 postula que *as órbitas de todos os corpos celestes são elipses*, ao passo que h_2 afirma simplesmente que *as órbitas de todos os planetas são elipses*.

Alguns esclarecimentos são necessários. Primeiro, Swinburne (2001, p. 82) admite que não há um modo acurado de mensurar o escopo de uma hipótese. Com-

[88] A demonstração pode ser feita pelo teorema 7 disponível no apêndice.

parações bastante gerais e vagas podem ser feitas entre as alegações de hipóteses competidoras. Segundo, pode acontecer que a probabilidade posterior de uma hipótese de maior escopo seja maior do que as probabilidades posteriores de hipóteses de menor escopo. Considere a teoria da Relatividade Geral de Einstein. Ela faz mais alegações sobre objetos e entidades, explica um maior número de fenômenos e realiza predições mais precisas do que a mecânica clássica de Newton. Embora ela tenha maior escopo do que a teoria de Newton, dada tais condições, ela é provavelmente mais verdadeira do que a de Newton. Ou seja, em comparação com a última, a teoria de Einstein satisfaz bem os critérios de encaixe com a evidência de fundo e de poder explanatório, mas não tão bem o de escopo. Por assim dizer, os outros dois critérios a compensam e aumentam a sua probabilidade posterior. Por conseguinte, uma hipótese h_1 que se encaixa bem com a evidência de fundo e tem grande poder explanatório, a despeito de ter escopo amplo, pode ter probabilidade posterior maior do que uma hipótese alternativa h_2 de menor escopo, especialmente quando h_2 não é bem-sucedida em relação aos outros dois critérios prévios. Terceiro, Swinburne (2001, p. 82 e 2004, p. 56) não considera esse critério tão fundamental quanto o critério de simplicidade. Por vezes, o escopo de uma hipótese é medido arbitrariamente e a hipótese perde simplicidade à medida em que o escopo é menor.

Todavia, considerando as probabilidades iniciais de duas hipóteses h_1 e h_2, se h_1 é uma hipótese de maior escopo do que h_2 e h_1 acarreta h_2, mas h_2 não acarreta h_1, então a probabilidade de h_1 pode ser no máximo igual[89] a de h_2. Serão iguais quando $Pr(\neg h_1 \wedge h_2) = 0$. Se, por outro lado, $Pr(\neg h_1 \wedge h_2) > 0$, a probabilidade h_1 será menor do que a de h_2. Omitindo a variável k, se $h_1 \vDash h_2$, então $Pr(h_1) \leq Pr(h_2)$.

4.4.4 Simplicidade

O critério de simplicidade é o mais determinante critério de probabilidade lógica de Swinburne.[90] Mais simples é uma hipótese, maior é a sua probabilidade inicial e, outras coisas sendo iguais, maior é a sua probabilidade posterior em comparação a

[89]Ver teorema 5.
[90]Sobre outras posições envolvidas no debate sobre simplicidade em Alan Baker (2010).

qualquer outra hipótese concorrente. Como dissemos, simplicidade é considerada um critério *a priori* na sua teoria. Se k é meramente tautológico, assumindo que h e k se encaixam apropriadamente, a probabilidade inicial de h, $Pr(h \mid k)$, dependerá, além do critério de escopo, de fatores intrínsecos a h. Neste caso, a probabilidade inicial de h é determinada de modo independente de qualquer evidência empírica. Tais fatores intrínsecos contemplam um conjunto de aspectos que constitui, em última instância, o conceito de simplicidade. Supondo uma partição de hipóteses $\Omega = \{h_1, h_2, \cdots, h_n\}$ onde todas elas se encaixam igualmente bem com a evidência de fundo, têm o mesmo escopo e predizem um conjunto amplo de evidências com grande poder explanatório, h_i tem maior probabilidade posterior somente se h_i é a mais simples de Ω. Essa é uma das principais teses de Swinburne (1997, p. 56 e 2001, p. 97). Mas, afinal, quais são tais características intrínsecas ao critério de simplicidade?

Os dois primeiros aspectos são simplicidades quantitativa e qualitativa.[91] O primeiro sentido de simplicidade corresponde ao número de instâncias ou ocorrências (*token*) de um mesmo tipo (*type*) de entidade (ou propriedade) que uma hipótese postula como existente. Assim, h_1 é mais simples quantitativamente do que h_2 se e somente se h_1 postula menos instâncias ou ocorrências de um mesmo tipo de entidade ou propriedade do que h_2. Por exemplo, h_1 afirma que todos os seres humanos são dotados de espírito, ao passo que h_2 sustenta que todos os seres vivos são dotados de espírito. O segundo sentido é relativo ao número de tipos de entidade ou propriedades que uma hipótese postula como existente. Dessa maneira, h_1 é mais simples qualitativamente do que h_2 se e somente se h_1 postula menos tipos de entidades ou propriedades do que h_2. Por exemplo, h_1 postula que existem meramente dois tipos distintos de *quark*, enquanto que h_2 postula que existem seis tipos diferentes de *quark*. De acordo com Swinburne (1997, p. 24 e 2001, p. 87), esses são os dois primeiros sentidos pelos quais uma hipótese é mais simples do que uma outra hipótese competidora.

Uma hipótese h_1 é mais simples do que uma hipótese h_2 se h_1 postula pro-

[91] Em *Counterfactuals* (1973, cap. 4, p. 87), David Lewis traça uma distinção entre esses dois sentidos de simplicidade, alegando que o seu realismo sobre mundos possíveis é apenas não simples ou não parcimonioso quantitativamente, mas é simples ou parcimonioso qualitativamente. Essa distinção corresponde exatamente aos dois primeiros aspectos de simplicidade de Swinburne.

priedades mais prontamente ou mais facilmente observáveis do que h_2. Esse terceiro aspecto do critério de simplicidade de Swinburne pode ser explicado nos seguintes termos. Uma propriedade D é mais prontamente observável do que uma propriedade F quando, para qualquer x, podemos descobrir se x é ou não D sem precisarmos descobrir se x é ou não F, mas não o contrário. Por exemplo, suponha o famoso caso da propriedade 'verdul' (*grue*). Para qualquer x, x é verdul em t se e somente se x é verde e t é anterior a 2050 ou x é azul e t é posterior a 2050. Assim, todos os objetos verdes observados até agora são verdul. Mas Swinburne (2004, p. 54) destaca que qualquer objeto x pode ser observado com a propriedade 'verde' sem precisarmos descobrir se t é anterior ou posterior a 2050 e sem precisarmos descobrir se x é 'verdul'. Nessa perspectiva, esse aspecto do critério revela que a propriedade 'verde' é mais prontamente observável do que a propriedade 'verdul' e uma hipótese que afirma que 'todas as esmeraldas são verdes' é mais simples do que uma que postula que 'todas as esmeraldas são verdul'.[92] Entretanto, como o próprio Swinburne (2004, p. 54) admite, algumas teorias científicas postulam propriedades não prontamente observáveis, como as noções de hipercarga e isospin na física de partículas elementares, mas são compensadas em relação aos outros critérios. Porque elas satisfazem muito bem o critério de poder explanatório e têm grande capacidade de explicar e predizer uma diversidade de fenômenos, elas podem ser mais prováveis do que outras hipóteses que postulam propriedades mais prontamente observáveis.

Segundo Swinburne (2001, p. 89-90), o quarto aspecto concerne ao número de leis independentes que uma hipótese postula. Outras coisas se mantendo constantes, uma hipótese h_1 que postula menos leis individuais do que h_2 é mais simples e tem maior probabilidade inicial. Quanto a esse aspecto, a teoria de Kepler sobre as três leis do movimento planetário é mais simples do que a teoria sobre o modelo geocêntrico de Ptolomeu. O quinto aspecto de simplicidade está intimamente conec-

[92]No contexto do novo problema ou enigma da indução, Nelson Goodman (1983 [1955]) concebeu o problema da propriedade 'verdul' e dos predicados projetáveis. O problema é que o argumento 'todas as esmeraldas observadas são verdul e, portanto, todas as esmeraldas são verdul' instancia o mesmo esquema inferencial indutivo de 'todas as esmeraldas observadas até agora são verdes, portanto, todas as esmeraldas são verdes', mas enquanto o segundo parece ser um bom argumento indutivo, o primeiro não. Observe que não é isso o que Swinburne está discutindo, mas sim o que é para uma propriedade ser prontamente observável.

tado com o quarto aspecto. Uma hipótese com poucas leis individuais relacionando poucas variáveis é preferível a hipóteses com muitas leis individuais relacionando muitas variáveis e a hipóteses com o mesmo número de leis relacionando mais variáveis do que a primeira. Por essa razão, para Swinburne (2001, p. 90), uma hipótese h_1 que postula duas leis individuais L_1 e L_2, onde L_1 relaciona as variáveis z e w e L_2 relaciona y e v, é mais simples do que uma hipótese h_2 que postula quatro leis L_1, L_2, L_3 e L_4 distintas e ainda relaciona um número maior de variáveis em cada uma dessas leis.

O sexto e último aspecto diz respeito a hipóteses matematicamente mais simples. Swinburne (1997, p. 26-27 e 2001, p. 90) distingue duas características internas envolvidas em tal aspecto. Por um lado, poucos termos em uma descrição ou equação matemática é preferível do ponto de vista da simplicidade. Por exemplo, $y = z + x$ relaciona menos termos do que $y = z + x + x^2$. Por isso, a primeira é mais simples matematicamente do que a segunda. Por outro lado, uma equação que descreve um conjunto de estados de coisas com um número menor de relações e entidades matemáticas é mais simples do que uma que descreve um número maior de relações e entidades. Swinburne (2001, p. 90) alega que se podemos aprender o significado de uma entidade (ou relação) matemática ϕ sem qualquer entendimento de outra ψ, mas não o inverso, então ϕ é mais simples do que ψ. Portanto, qualquer número inteiro é mais simples do que qualquer número racional e este, por sua vez, é mais simples do que qualquer número real ($\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$), adição é uma relação mais simples do que multiplicação, e assim por diante. Usando o seu exemplo, $y = x$ é matematicamente mais simples do que $y = \sqrt{5x}$.

Swinburne (2001, p. 83-84) oferece um caso no qual o critério de simplicidade desempenha papel crucial na escolha entre duas hipóteses de igual escopo, que parecem ter o mesmo grau de probabilidade em relação ao critério de poder explanatório e se encaixam adequadamente bem com a evidência de fundo. Em primeiro lugar, suponha que duas variáveis x e y estejam sendo estudadas em uma área de investigação científica, mas sem qualquer evidência ou conhecimento de fundo relevante para decidir entre duas hipóteses concorrentes. Assim, a tabela abaixo mostra o que se tem encontrado de observação sobre x e y. É, por assim dizer, a evidência disponível que se tem até o momento:

x	0	1	2	3	4	5	6
y	0	2	4	6	8	10	12

Assim, temos uma hipótese h_1 que postula a seguinte equação como explicação da evidência: $y = 2x$. Com essa fórmula, é possível inferir os valores observados em y a partir dos valores correspondentes observados em x. Ou seja, um *input* em x gera um *output* correspondente em y e a equação $y = 2x$ é capaz de explicar muito bem tal correspondência. Porém, outra hipótese, vamos chamar de h_2, é igualmente capaz de explicar tal evidência disponível: $y = 2x + x.(x-1).(x-2).(x-3).(x-4).(x-5).(x-6).z$. Na verdade, infinitas fórmulas dessa última forma podem gerar um valor para y a partir dos valores de x, onde z pode ser uma constante ou alguma função de x. Em tais situações, Swinburne (2001, p. 83-84) defende que devemos escolher pela mais simples, a saber, $y = 2x$. Para Swinburne, ela é provavelmente a verdadeira. A propósito, a última equação é redutível à primeira se $z = 0$. Podemos também predizer valores para y adicionando valores em x. Por exemplo, se $x = 7$, então $y = 14$; se $x = 8$, então $y = 16$; $y = 18$ se $x = 9$ e assim por diante. Teríamos, obviamente, que verificar se as predições são de fato corretas. Apesar disso, a questão é que ambas hipóteses produzem o mesmo *output* a partir de um determinado *input*, considerando que ambas satisfazem bem todos os outros critérios. Afinal, a hipótese h_1 é mais provável do que h_2 por explicar a evidência com uma equação matematicamente mais simples? Em tal cenário, podemos reservar dúvidas se esse é um critério estritamente epistêmico, um que seja conducente à verdade, ou se é somente um critério estético.

Alguns comentários finais são dignos de consideração. Primeiro, ainda que a maioria dos exemplos de Swinburne sejam respectivos a hipóteses e teorias de natureza científica, todos os seus critérios se aplicam igualmente aos dois padrões de explicação, científica e pessoal. Segundo, uma hipótese pode ter um conjunto de formulações logicamente equivalentes. Naturalmente, Swinburne (2001, p. 87) prefere a mais simples delas. Se F é a formulação mais simples de uma hipótese

h e F' é a mais simples de h', então h é mais simples do que h' se e somente se F é mais simples do que F'. Terceiro, Swinburne (1997, p. 29 e 2001, p. 91) reconhece que hipóteses podem perder algum aspecto de simplicidade dependendo da sua formulação. É possível reduzir as equações de Maxwell, que foram originalmente propostas com quatro leis e na forma de vetores e grandezas escalares, em termos de tensor eletromagnético, mas com apenas duas leis gerais. Embora a segunda tenha menos leis individuais, a primeira tem menos variáveis e o conceito de tensor eletromagnético pressupõe entendimento dos conceitos de campo elétrico e campo magnético. Ganha-se simplicidade em relação a um aspecto, mas perde-se simplicidade no tocante a outro aspecto. Quarto, Swinburne (2001, p. 85-86) deixa claro que a sua concepção de simplicidade é distinta da proposta de Popper (2002 [1959]). Este último identifica simplicidade de uma hipótese ou teoria com o seu escopo — hipóteses mais simples têm maior escopo — e com graus de falseabilidade. Mas esse sentido de simplicidade é rejeitado por Swinburne. Quinto, a estratégia de Swinburne está clara (1997, p. 56 e 2001, p. 102-107). Uma hipótese mais simples h_1 tem maior probabilidade inicial do que as suas concorrentes $h_2, h_3, \cdots, h_n$ e, quando todas elas satisfazem bem todos os outros três critérios (encaixe com a evidência de fundo, poder explanatório e escopo), a sua probabilidade posterior será maior do que qualquer outra hipótese rival.[93] Se k é meramente tautológico, ou na ausência de qualquer evidência de fundo relevante, a alegação é de que há um modo objetivo de determinar se a probabilidade lógica inicial de uma hipótese é maior do que a de outra, especialmente pelos fatores descritos no critério de simplicidade. Sexto, colocando em termos mais técnicos, supondo uma partição de teorias ou hipóteses $\Omega = \{h_1, h_2, \cdots, h_n\}$, se h_1 satisfaz melhor todos os aspectos do critério de simplicidade do que qualquer outra hipótese h_i de Ω ($i \neq 1$), então h_1 é a mais simples. O problema, no entanto, é que tais comparações não são muito precisas. Uma hipótese h_1 pode satisfazer melhor um aspecto do que h_2, mas outra hipótese h_3 pode satisfazer melhor outros dois aspectos do que h_2 e assim por diante. Além disso, não

[93]'[...]Greater simplicity means greater prior probability, and so — for given e and non-empirical k — greater posterior probability. Bayes's theorem allows us to give formal articulation to this claim' (1997, p. 56).

é exatamente claro como uma avaliação global de tais fatores deve ser conduzida. Ou seja, se a ordem que Swinburne oferece é meramente arbitrária ou se espelha uma hierarquia ou, ainda, se há uma hierarquia alternativa desses aspectos de simplicidade. Para o conhecimento de fundo k e uma mesma evidência e, assumindo que $h_1, h_2, \cdots, h_n$ cumprem igualmente bem os outros três critérios, torna-se difícil determinar que a probabilidade posterior da hipótese mais simples h_1 é maior do que as probabilidades posteriores das suas rivais se não temos como medir o impacto dos fatores de simplicidade sobre a probabilidade inicial de cada hipótese e se eles realmente aumentam a probabilidade inicial de h_1; ou seja, se simplicidade significa de fato maior probabilidade inicial e, outras coisas sendo iguais, maior probabilidade posterior. Em todo caso, vamos explorar algumas consequências, problemas e objeções na última seção.

4.5 Princípio de Indiferença

Considere um mecanismo de jogadas de um determinado *dado*. O seu espaço de possibilidades, ou espaço de amostra, é constituído de seis resultados: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Vamos supor que não temos nenhuma evidência ou informação de fundo relevante sobre o mecanismo de jogadas, se é ou não justo, se o dado é fisicamente simétrico e, ademais, somos ignorantes se há uma tendência maior em sair um resultado do que outro. No entanto, vamos assumir que temos unicamente informações ou evidências sobre os resultados possíveis, isto é, temos boas razões para pensar que a configuração do espaço de possibilidades é composta pelos resultados em Ω. Devemos distribuir de modo simétrico os graus de probabilidade inicial (*priors*) entre os seis resultados possíveis? Em outros termos, uma vez que não temos mais informações, é racional atribuir probabilidade de $\frac{1}{6}$ a cada um dos resultados desse espaço de possibilidades? Ou qualquer outra atribuição de probabilidade situando-se entre $[0, 1]$ seria igualmente correta?

Esse problema é matéria de disputa entre Bayesianos objetivos e Bayesianos subjetivos. Mais precisamente, o que está em discussão aqui é o princípio de indiferença. Alegadamente, esse princípio diz que se um agente S tem unicamente a evidência ou informação de que um *espaço de amostra* Ω tem n possibilidades mutuamente exclusivas e conjuntamente exaustivas (uma partição), $\Omega = \{h_1, h_2, \cdots, h_n\}$, então é racional para S atribuir grau de $\frac{1}{n}$ a cada uma de tais possibilidades de Ω. Esse princípio funciona como um critério *a priori* de atribuição de probabilidades iniciais. Mas enquanto Bayesianos objetivos estão dispostos a aceitar princípios de tal natureza, Bayesianos subjetivos, em sua versão mais radical, rejeitam sumamente critérios *a priori* sobre as probabilidades iniciais.[94] Estes não exigem nenhuma outra condição restritiva adicional, além da conformidade com o cálculo de probabilidades (ou seja, graus coerentes probabilisticamente).

Entretanto, voltando ao nosso exemplo, podemos dividir o espaço de possibilidades de diversas maneiras. Em vez de $\Omega = \{1, 2, 3, 4, 5, 6\}$, poderíamos conceber um espaço de possibilidades alternativo tal que $\Omega' = \{obter\ 3, \neg\ obter\ 3\}$ ou tal que $\Omega'' = \{\text{par, ímpar}\}$. A atribuição de probabilidade inicial de obter o resultado de 3, $Pr(obter\ 3)$, não é de $\frac{1}{6}$, mas de $\frac{1}{2}$ se consideramos Ω' como a partição correta. Consequentemente, descrições diferentes de um mesmo espaço de possibilidades podem resultar em atribuições incoerentes de probabilidades iniciais. Mas qual é a atribuição correta? Por que privilegiar uma divisão do espaço de possibilidades sobre outra? Afinal, qual é o modo correto de aplicar o princípio de indiferença?

Essa é uma versão simples do paradoxo de Bertrand[95] (2011 [1888]). A versão que Swinburne (2001, p. 116-117) discute é mais sofisticada e podemos apresentá-la como se segue. Considere um parâmetro T do qual temos a única informação de que o seu valor se situa em um intervalo $[a, b]$. Podemos atribuir graus simétricos de probabilidade inicial de que o valor de T está entre a e $(a + 1)$ ou entre $(a + 1)$ e $(a + 2)$. Agora, considere que o valor de um parâmetro alternativo T' se situa entre a^2 e b^2 e essa é a única informação que possuímos. Pelo princípio de indiferença, o grau de probabilidade inicial de que o valor está entre a^2 e $(a^2 + 1)$ ou está entre $(a^2 + 1)$ e $(a^2 + 2)$ deve ser simetricamente distribuído. Ainda, o valor de T' pode ser

[94] O princípio de indiferença foi originalmente concebido por Laplace (1951 [1820]), mas foi J. M. Keynes (1921) quem cunhou a expressão. Mais informações sobre diferentes versões do princípio de indiferença em Colin Howson (2009), Jon Williamson (2010, cap. 2) e Jonathan Weisberg (2015, sec. 2.1). Um resumo da disputa entre Bayesianos em William Talbott (2008, sec. 4.2).

[95] Darren Bradley (2015, cap. 5, p. 69-73) e Jonathan Weisberg (2015, sec. 2.1) apresentam uma boa introdução a esse problema.

distribuído simetricamente no mesmo espaço de possibilidades, mas em um outro intervalo, ou seja, entre a^2 e $a^2 + (2a + 1)$ ou entre $a^2 + (2a + 1)$ e $a^2 + (4a + 2)$. A questão novamente é que podemos ter graus incoerentes de probabilidade inicial para um mesmo espaço de possibilidades. Swinburne (2001, p. 117) defende que o princípio de indiferença deve ser usado na sua formulação mais simples e para possibilidades de igual escopo. No seu exemplo, a formulação de que T se situa entre $[a, b]$, tão provável que esteja entre a e $(a + 1)$ ou entre $(a + 1)$ e $(a + 2)$, é mais simples do que as formulações sobre o valor do parâmetro T'. Em tais casos onde não temos nenhuma outra informação privilegiada sobre a partição, exceto a informação de que ela é constiuída de n possibilidades ou um valor ε se situa em um determinado intervalo $[a, b]$, Swinburne (2001, p. 118) se questiona como seria possível realizar qualquer atribuição de probabilidade se rejeitássemos atribuições de probabilidades iniciais baseadas em critérios *a priori*. Simplesmente alegar que qualquer atribuição entre $[0, 1]$ é racional, desde que seja coerente probabilisticamente, parece insuficiente. Contudo, não é claro como os fatores do critério de simplicidade devem ser aplicados sobre o princípio de indiferença de modo a privilegiar um espaço de possibilidades sobre outros. Além de não esclarecer esse ponto, torna mais confuso um problema que constitui desafio a Bayesianos objetivos.

4.6 Confirmação Absoluta e Confirmação Incremental

Pelo teorema de Bayes, podemos medir o quão bem uma hipótese é suportada evidencialmente. Tipicamente, a teoria Bayesiana de confirmação estabelece uma metodologia a partir da qual podemos mensurar probabilisticamente relações de confirmação, desconfirmação e neutralidade ou irrelevância evidencial entre um conjunto de evidências e uma hipótese. Assim, se e suporta evidencialmente h, embora em um grau não conclusivo, então e confirma h. Se e refuta h, então e desconfirma h. Se e não suporta e nem refuta h, então e é evidencialmente irrelevante ou neutra em relação à hipótese h. Aqui, novamente, estamos considerando a evidência total; não apenas uma nova peça de evidência e, mas igualmente toda evidência ou conhecimento de fundo k. Portanto, relações de confirmação, desconfirmação ou irrelevância evidencial que e fornece a h devem ser relativas a k. À medida em que novas peças de evidências $e_1, e_2, \cdots, e_n$ são obtidas, elas devem ser condicionalizadas e incorporadas ao estoque de evidência total k.

Podemos, no entanto, distinguir entre dois conceitos de confirmação. Quando uma hipótese é confirmada em sentido absoluto por um conjunto de evidências, significa que o seu grau de suporte evidencial probabilístico supera um determinado limiar de valor χ (*threshold*) apropriado. Nesse sentido, h é fortemente suportada por e. Definimos confirmação absoluta[96] nos seguintes termos: considerando k, e confirma h se e somente se $Pr(h \mid e \wedge k) > \chi$. Normalmente, é estipulado que $\chi = .5$. Se $Pr(h \mid e \wedge k) > .5$, então $Pr(\neg h \mid e \wedge k) < .5$. Mas nem sempre é possível determinar de maneira acurada o grau de suporte evidencial com valores numéricos precisos e podemos questionar se $\frac{1}{2}$ é ou não um valor arbitrariamente fixado para confirmação absoluta.

Por essas razões, parece mais adequado adotar o conceito de confirmação incremental[97] em substituição ao conceito de confirmação absoluta. Desconfirmação, irrelevância e confirmação podem ser definidas como se segue:

- **e** desconfirma **h** *sse* $Pr(h \mid e \wedge k) < Pr(h \mid k)$
- **e** é evidencialmente irrelevante para **h** *sse* $Pr(h \mid e \wedge k) = Pr(h \mid k)$
- **e** confirma **h** *sse* $Pr(h \mid e \wedge k) > Pr(h \mid k)$

Observe que e pode confirmar incrementalmente h sem confirmá-la absolutamente. Não se segue de $Pr(h \mid e \wedge k) > Pr(h \mid k)$ que $Pr(h \mid e \wedge k) > .5$, não necessariamente. Dado k, se e incrementalmente confirma h, então e aumenta a probabilidade de h, mas pode ser um valor inferior a .5. O suporte evidencial que e fornece a h é suficiente para tornar h mais provável do que simplesmente a probabilidade de h condicional em k. Dito de outra maneira, se e oferece confirmação incremental a h, significa que h é menos provável na ausência da evidência e, a saber, o impacto de e sobre h incrementa o valor de probabilidade de h condicional somente em k. Avalia-se, neste caso, a força que e desempenha sobre h e a força de h sem e

[96] Ver John Earman (1992, cap. 3, p. 66-67) e Jonathan Weisberg (2011, p. 535).
[97] Ver James Joyce (2004, p. 143) e Howson e Urbach (2006, cap. 4, p. 92).

em termos comparativos. Além disso, a comparação entre $Pr(h \mid k)$ e $Pr(h \mid e \wedge k)$ nos proporciona uma medida de confirmação, a medida de diferença d:

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

Portanto, $d(h, e, k) < 0$ se $Pr(h \mid e \wedge k) < Pr(h \mid k)$, $d(h, e, k) = 0$ se $Pr(h \mid e \wedge k) = Pr(h \mid k)$ e $d(h, e, k) > 0$ se $Pr(h \mid e \wedge k) > Pr(h \mid k)$. Em resumo, desconfirmação se < 0, neutralidade evidencial se $= 0$ e confirmação se > 0.

Vale dizer que a teoria Bayesiana oferece um amplo conjunto de medidas alternativas de confirmação, cada uma capturando um fator diferente de impacto de e sobre h.[98] Para os nossos propósitos atuais, entretanto, a medida de diferença d, a medida clássica, será avaliada na última seção. Ela apresenta uma séria deficiência no tocante ao seu poder de confirmação, especificamente se certas condições são satisfeitas; talvez o desafio mais contundente ao Bayesianismo seja o problema da evidência antiga. De qualquer forma, Swinburne (2001, p. 104 e 2004, p. 17) está disposto não somente aceitar os dois conceitos de confirmação, também os empregando na sua classificação de argumentos indutivos, mas endossa a medida d.

Por último, como dissemos, Swinburne (2005, p. 17) aplica os conceitos de confirmação a certos padrões de inferência indutiva. Ele entende que existem dois tipos de argumentos indutivos corretos, argumentos P-indutivos e C-indutivos, cada um correspondendo a um conceito distinto de confirmação. Por um lado, supondo que e representa um conjunto de premissas de um argumento em suporte de uma conclusão h, quando e confirma absolutamente h, temos um argumento P-indutivo correto. Por outro lado, quando e confirma incrementalmente h, temos um argumento C-indutivo correto. Ambos sob a suposição de um conhecimento de fundo k. Por exemplo, '80% dos habitantes de Porto Alegre são católicos' e 'João é habitante de Porto Alegre' torna 'João é católico' provável no primeiro

[98] Existe uma diversidade de medidas de confirmação. Para citar algumas delas, a medida s, $Pr(h \mid e \wedge k) - Pr(h \mid \neg e \wedge k)$, defendida por David Christensen (1999) e James Joyce (2004), 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*, $\log[\frac{Pr(e|h \wedge k)}{Pr(e|\neg h \wedge k)}]$, medida l, defendida por Branden Fitelson (1999). Outras medidas disponíveis em Franz Huber (2010, sec. 6). Seguindo John Earman (1992, cap. 5, p. 121), falamos brevemente por que razão a medida s parece inapropriada em André Neiva (2015).

sentido; 'todos os corvos observados até agora são pretos' torna 'todos os corvos são pretos' provável no segundo sentido. Ou seja, o primeiro tipo de argumento corresponde a $Pr(h \mid e \wedge k) > \chi$, onde $\chi = .5$, e o segundo tipo diz respeito a $Pr(h \mid e \wedge k) > Pr(h \mid k)$.

4.7 Problemas e Objeções

Embora tenhamos discutido algumas críticas ao longo deste trabalho, vamos nos voltar mais detidamente a alguns dos principais problemas e objeções que podem ser levantados contra a teoria de Swinburne e ao Bayesianismo em geral.

Primeiro, as definições de probabilidade indutiva de Swinburne são relativas a agentes epistêmicos com capacidades lógicas distintas, desde uma mais restrita a um modelo mais ideal. Agentes ordinários nem sempre são capazes de realizar inferências de modo competente, de atribuir graus corretos de probabilidade indutiva, de usar os padrões e critérios de probabilidade lógica, e assim por diante. A suposição é de que agentes logicamente oniscientes são capazes de atribuir probabilidade 1 a tautologias e 0 a contradições, a tirar consequências dedutivas dos axiomas do *calculus*, entre outras *performances* consideradas ideais. Como insistimos em vários momentos, o ônus dessa proposta é se comprometer com um modelo de perfeição racional que não parece condizente com práticas mais ordinárias e, mais gravemente, parece ser um modelo inatingível. A propósito, a objeção da onisciência lógica[99] se alastra a toda teoria que reivindica que o cálculo probabilístico deve ser satisfeito por agentes epistêmicos. Mais especificamente, como dissemos em outra oportunidade, o axioma de normalização ou certeza exige que agentes atribuam grau máximo de probabilidade a tautologias em geral, não somente as conhecidas. Isso é de fato um requerimento muito exigente, pois nos faltam as razões corretas para crer justificadamente e atribuir grau de probabilidade 1 a muitos teoremas e verdades lógicas. Todavia, uma possível resposta a essa objeção poderia ser articulada nos seguintes termos. Ora, os padrões exigidos pelo Bayesianismo, e particularmente pela

[99]Mais sobre a suposição da onisciência lógica como exigência imposta pelo Bayesianismo em Daniel Garber (1983) e William Talbott (2008, sec. 6).

teoria de Swinburne, são normativos; modelos para agentes idealmente racionais e logicamente oniscientes. Se falhamos em seguir e satisfazer tais padrões, falta de sorte nossa. Em sentido rigoroso, a teoria impõe as condições e o modo correto pelos quais atribuições de probabilidade indutiva devem ser realizadas.

Segundo, Swinburne (2001, p. 121 e 2002, p. 8) admite que agentes devem satisfazer o cálculo, embora isso seja uma restrição mínima. Defendendo uma proposta de Bayesianismo objetivo, a motivação é de que existe uma maneira objetiva de restringir as probabilidades iniciais, a saber, o cálculo não nos fornece isso porque ele é incompleto. Destarte, os critérios de escopo e simplicidade e o princípio de indiferença desempenham função de restrições adicionais na sua teoria. Em todo caso, Swinburne não oferece um argumento a favor da tese de que é racional para agentes obedecerem ao maquinário de probabilidades, embora ele concorde com a condição de coerência probabilística. Como vimos, o probabilismo concerne a uma parte do programa de epistemologia Bayesiana, a sua dimensão sincrônica, e defende duas teses: (*i*) graus de crença e (*ii*) a norma de racionalidade e coerência probabilística. O argumento do *Dutch Book* pretende defender, sobretudo, a tese (*ii*) e é controverso principalmente por causa da premissa sobre a correspondência entre quocientes de aposta e graus de probabilidade subjetiva e porque, em última instância, consiste em uma defesa pragmática. Em outras palavras, o *Dutch Book* apela a comportamentos de apostas e se ocupa mais diretamente da dimensão de racionalidade pragmática. Embora Swinburne não aceite a tese (*i*), uma vez que advoga o modelo de crença *simpliciter*, existem outros tipos de argumento em defesa da tese (*ii*) do probabilismo que nem sequer são avaliados nas suas obras. Nessa esteira, Patrick Maher (1993 e 1997) oferece um argumento baseado no teorema representacional e na noção de preferências, onde estas precisam obedecer a certas restrições básicas de racionalidade. O teorema representacional pretende unificar utilidades e probabilidades subjetivas de agentes numa função única de utilidade esperada. Alternativamente, como mencionamos em uma nota, o argumento da acurácia de James Joyce (1998 e 2009) se ocupa de fins epistêmicos e pretende demonstrar que funções probabilísticas de graus de crença são mais acuradas e não são dominadas por nenhuma outra função não-probabilística alternativa. Mais recentemente, Richard Pettigrew (2013, 2015 e 2016 (*forthcoming*)) tem sustentado um programa de *accuracy-first epistemology*, em muitos aspectos influenciado pela proposta de Joyce (1998), com o desenvolvimento da teoria de utilidade epistêmica. Esses argumentos podem ser questionados, mas são defesas mais promissoras do que o clássico argumento do *Dutch Book* ou contrato de perda garantida, além de promoverem uma discussão fecunda e nova na agenda contemporânea em epistemologia formal. Swinburne (2001, p. 119-123) reconhece algumas deficiências do *Dutch Book*, mas, em contrapartida, não oferece nenhum argumento em prol da tese sobre a norma de coerência probabilística.

Terceiro, o critério de simplicidade de Swinburne funciona como um princípio *a priori* sobre probabilidades iniciais de hipóteses. Quando estas estão condicionadas em evidência ou conhecimento tautológico k, supostamente o critério de escopo e os fatores intrínsecos de simplicidade determinam o valor de probabilidade da função $Pr(h \mid k)$ de modo independente de qualquer evidência empírica. Alega-se que tais critérios permitem determinar objetivamente se uma hipótese tem maior probabilidade inicial do que outras. Essa é uma afirmação importante de Swinburne. No entanto, ainda que não seja necessário que essa função tenha valores numéricos exatos, não está claro como os próprios aspectos e fatores do critério de simplicidade aumentam o grau de probabilidade inicial de uma hipótese. Swinburne não apresenta nenhum modelo formal ou explicação mais precisa de como esses seis fatores de simplicidade devem ser computados na função $Pr(h \mid k)$; se algum deles prevalece sobre os outros ou se há algum meta-critério no qual uma ordem com relações de subordinação e relevância pode ser estabelecida. Alega-se que uma hipótese mais simples terá probabilidade maior do que as suas concorrentes. Entretanto, como dissemos, pode ser o caso que uma hipótese h_1 satisfaça melhor o aspecto de ser matematicamente simples, que uma hipótese h_2 corresponda melhor aos fatores de simplicidade quantitativa e qualitativa, que h_3 melhor satisfaça esses três fatores do que h_1 e h_2, mas não corresponda bem a um outro aspecto qualquer, e assim sucessivamente com comparações entre inúmeras hipóteses. Em termos de *otimização*, melhor uma hipótese satisfaz todos os fatores, mais simples ela é e, supostamente, maior é a sua probabilidade inicial. Mas nem sempre hipóteses conseguem satisfazer completamente todos os seis aspectos de simplicidade. Mais uma vez, nenhuma hierarquia é estipulada e nenhuma medida de como esses fatores incrementam o grau de probabilidade na função $Pr(h \mid k)$ é oferecida. Considerando uma mesma evidência e, o conhecimento tautológico k e uma partição $\Omega = \{h_1, h_2, \cdots, h_n\}$ onde h_i é a mais simples das alternativas, se não conseguimos determinar como os fatores de simplicidade aumentam a probabilidade inicial de h_i, não é possível determinar se a probabilidade posterior $Pr(h_i \mid e \wedge k)$, o *output* gerado pelo teorema de Bayes, será maior do que as probabilidades posteriores das outras hipóteses da partição, supondo que todas h_1, h_2, $\cdots$, h_n obedecem igualmente bem aos outros critérios de probabilidade lógica. Para maior clareza da exposição, vamos considerar duas hipóteses mutuamente exclusivas e conjuntamente exaustivas h_1 e h_2. Ambas satisfazem bem todos os outros critérios e h_1 é mais simples do que h_2. Como colocamos anteriormente, o teorema de Bayes nos permite a seguinte comparação: $Pr(e \mid k) \neq 0$, se $Pr(e \mid h_1 \wedge k) = Pr(e \mid h_2 \wedge k)$, então $Pr(h_1 \mid e \wedge k) > Pr(h_2 \mid e \wedge k)$ se e somente se $Pr(h_1 \mid k) > Pr(h_2 \mid k)$. Outras coisas sendo iguais, a probabilidade posterior de h_1 será maior do que a de h_2 se e somente se a probabilidade inicial de h_1 for maior do que a de h_2. Mas por que o princípio de simplicidade deve ser o critério decisivo? E como ele incrementa a probabilidade inicial de h_1 de tal sorte que a torna maior do que a da sua rival h_2?[100]

Quarto, o problema da evidência antiga é o que John Earman (1992, p. 119) chamou de o *calcanhar de Aquiles* do Bayesianismo.[101] Vamos admitir que em um tempo t uma evidência antiga e suporta uma hipótese h, dado k, com um grau $0 < \chi < 1$. Suponhamos que em t' descobre-se uma relação lógica tal que a conjunção $(h \wedge k)$ acarrete e. Portanto, $(h \wedge k) \vDash e$. Igualmente, vamos assumir que $Pr(e \mid k) = 1$ em t', ou seja, e é certa condicional em k. Se $(h \wedge k) \vDash e$, então $(h \wedge k) \equiv ((h \wedge k) \wedge e)$. Assim, se $(h \wedge k) \equiv ((h \wedge k) \wedge e)$, então $Pr(h \wedge k) = Pr(h \wedge k \wedge e)$. Pela definição de pro-

[100]Uma avaliação sobre o princípio de simplicidade como um critério epistêmico, outras abordagens desse critério no contexto Bayesiano e problemas envolvendo essas propostas está disponível em Elliott Sober (2015, cap. 2).

[101]Clark Glymour (1980, p. 85-93) originalmente formulou o problema da evidência antiga. Uma exposição útil e organizada de versões diferentes do problema da evidência antiga em Ellery Eells (1990, p. 207). Aqui apresentamos a versão *default*.

babilidade condicional, $Pr(e \mid h \wedge k) = \frac{Pr(e \wedge h \wedge k)}{Pr(h \wedge k)}$. Segue-se, por substituição, que $Pr(e \mid h \wedge k) = \frac{Pr(h \wedge k)}{Pr(h \wedge k)}$. Portanto, $Pr(e \mid h \wedge k) = 1$, assumindo que $Pr(h \wedge k) \neq 0$. Pelo teorema de Bayes, temos $Pr(h \mid e \wedge k) = \frac{Pr(e \mid h \wedge k) \times Pr(h|k)}{Pr(e|k)}$. Dado que $Pr(e \mid h \wedge k) = 1$ e $Pr(e \mid k) = 1$ em t', conclui-se que $Pr(h \mid e \wedge k) = Pr(h \mid k)$ em t'. Por consequência, assumindo o conceito de confirmação incremental e a medida de diferença $d(h, e, k) = Pr(h \mid e \wedge k) - Pr(h \mid k)$, depreende-se que $d(h, e, k) = 0$. Portanto, e, que era uma evidência antiga que suportava h em um certo grau em t, não confirma incrementalmente h em t'. Em outras palavras, sob a suposição das condições previamente colocadas, a medida de diferença d revela que e é evidencialmente irrelevante ou neutra em relação a h em t'. Aparentemente e perdeu o seu poder confirmatório sobre h.

Sobre esse último problema, algumas considerações merecem destaque. Primeira, se $Pr(e \mid k) = 1$ e $Pr(e \mid h \wedge k) = 1$, então $Pr(e \mid \neg h \wedge k) = 1$, considerando as três funções em t'. Supondo que $Pr(h \mid k) \neq 0$, pelo teorema de probabilidade total, $Pr(e \mid k) = [Pr(h \mid k) \times Pr(e \mid h \wedge k)] + [Pr(\neg h \mid k) \times Pr(e \mid \neg h \wedge k)]$. Assim, $[Pr(h \mid k) \times 1] + [Pr(\neg h \mid k) \times Pr(e \mid \neg h \wedge k)] = 1$ e, como corolário, $Pr(h \mid k) + [Pr(\neg h \mid k) \times Pr(e \mid \neg h \wedge k)] = 1$. Temos, por conseguinte, $Pr(\neg h \mid k) \times Pr(e \mid \neg h \wedge k) = 1 - Pr(h \mid k)$. Uma vez que $Pr(\neg h \mid k) = 1 - Pr(h \mid k)$, $Pr(\neg h \mid k) \times Pr(e \mid \neg h \wedge k) = Pr(\neg h \mid k)$. Portanto, $Pr(e \mid \neg h \wedge k) = 1$ se $Pr(\neg h \mid k) \neq 0$. Dessa maneira, h e $\neg h$ predizem igualmente bem e; os seus *likelihoods* e o seu poder explanatório têm grau máximo. Segunda, Swinburne não discute alguma solução ao problema. Como vimos, entretanto, é irrelevante para Swinburne se uma nova relação lógica entre e e h foi descoberta em um tempo t' ou mesmo se h foi formulada antes ou depois de e. Os graus corretos de probabilidade indutiva, o que ele entende por probabilidade lógica, não dependem de tais fatores. É uma questão objetiva se h e k acarretam e, $(h \wedge k) \vDash e$, e se $Pr(e \mid h \wedge k) = 1$; é independente do fato de agentes mais limitados serem ou não capazes de reconhecer essa relação. Nesse sentido, o poder de confirmação de e sobre h é o mesmo em t ou t', se estamos tratando de probabilidades lógicas. Talvez esse tipo de resposta seja a mais condizente com o que Swinburne propõe em sua teoria, embora possamos reservar dúvidas se ela efetivamente dissolve o problema. De qualquer forma, e esta é a terceira consideração, o problema da evidência antiga continua sendo uma crítica incisiva a versões de Bayesianismo que empregam o conceito de confirmação incremental e a medida de diferença d. Em última análise, ele nos revela que alguns reparos precisam ser realizados e medidas de confirmação mais plausíveis e apropriadas devem ser oferecidas como alternativas teóricas.[102]

[102] Talvez seja correto dizer que o conceito de confirmação incremental e a medida d constituam o que Jonathan Weisberg (2001, p. 535) chamou de uma *visão ingênua de confirmação Bayesiana*. Isso obriga Bayesianos a procurar por uma nova formatação de suas teorias e a propor soluções ao problema da evidência antiga. Mais recentemente, sobre esse problema, Branden Fitelson e Stephan Hartmann (2015) forneceram uma resposta bem articulada no estilo de Daniel Garber (1983).

5 Considerações Finais

No capítulo inaugural, tratamos de algumas relações entre crença e probabilidade na teoria de Swinburne. Vimos que sérias objeções podem ser formuladas contra a sua proposta de que crer que p é equivalente a crer que *p é mais provável do que $\neg p$*. Nem sempre crenças simples correspondem a crenças sobre probabilidades. As objeções de William Alston, sobretudo a do regresso vicioso *ad infinitum*, são cruciais e o próprio Swinburne as reconhece. Em seguida, exploramos algumas relações lógicas mais plausíveis entre crença *simpliciter* e crença sobre probabilidades. A despeito da análise de Swinburne ser apropriada, reiteramos que nem sempre agentes doxásticos ordinários têm crenças sobre probabilidades. Embora o modelo de graus de crença não seja canônico entre epistemólogos, ele fornece uma explicação mais plausível de casos onde um agente crê mais fortemente que p do que $\neg p$. Observamos em uma nota as alternativas de modelo formais para graus de crença: teoria de probabilidade, função de graus de crença Dempster-Shafer e medidas de possibilidade. Uma avaliação das diferenças entre tais modelos e das relações entre crença *simpliciter* e graus de crença, se abordagens unificacionistas são bem-sucedidas, constitui importante prospecto de pesquisa. Também tratamos brevemente de algumas propriedades de crença *simpliciter*, da distinção entre racionalidades epistêmica e pragmática e das implicações do conceito de crença de meios-a-fins para o âmbito da ação. A propósito, graus de crença e preferências estão intimamente conectadas em propostas Bayesianas em teoria da decisão, particularmente com o uso do teorema representacional como modelo padrão.

No capítulo intermediário, apresentamos o aparato formal do cálculo de probabilidades, as suas principais definições e analisamos algumas consequências que podem ser extraídas dele. Traçamos algumas breves distinções entre a axiomatização clássica e uma axiomatização alternativa no final da seção 3.2.3, a definição de probabilidade categórica é considerada primitiva na primeira e, diferentemente, o conceito de probabilidade condicional é tomado como primitivo na segunda. Esclarecemos e discutimos algumas interpretações e tipos de probabilidade. Na concepção de Swinburne, probabilidades físicas são explicadas com base em propensões naturais e probabilidades estatísticas são definidas em termos de frequência relativa, apesar de propensões e frequências serem tipicamente consideradas rivais na explicação do conceito de *chance*. A sua tipologia é de fato bastante peculiar nas suas subdivisões de probabilidade indutiva. Probabilidades subjetivas, epistêmicas e lógicas não são apenas relativas à evidência total, mas correspondem a agentes com capacidades lógicas distintas e referem-se a usos corretos ou incorretos dos critérios de probabilidade lógica. Por último, examinamos algumas aproximações entre probabilidade e justificação epistêmica. Salvo melhor juízo, probabilidades indutivas não parecem corresponder à justificação epistêmica. Em todo caso, reservamos dúvidas e cautela neste ponto, *sub judice* até uma investigação mais detalhada. Em última análise, probabilidade lógica é o tipo mais importante da teoria de Swinburne. O ônus da sua definição, no entanto, é associa-lá estritamente a graus corretos de probabilidade indutiva que agentes logicamente oniscientes são capazes de alcançar.

No último capítulo, consideramos em mais detalhes a proposta de Bayesianismo objetivo de Swinburne. Antes fizemos algumas distinções entre as posições de Bayesianos subjetivos e Bayesianos objetivos. Por um lado, Bayesianos subjetivos exigem que atribuições de probabilidade sejam coerentes probabilisticamente. Se um agente viola o *calculus*, então ele está vulnerável a um *Dutch Book* ou contrato de perda garantida. Expomos o argumento do *Dutch Book* e mostramos algumas das suas fraquezas. Por outro lado, fora a conformidade com o *calculus*, Bayesianos objetivos reivindicam princípios objetivos adicionais sobre os *priors*. Depois, analisamos e discutimos os critérios de probabilidade lógica de Swinburne. Os critérios de simplicidade e escopo são considerados princípios *a priori*, ao passo que encaixe com a evidência de fundo e poder explanatório são critérios *a posteriori*. Se k é meramente tautológico, então a função $Pr(h \mid k)$ é determinada principalmente pelos aspectos intrínsecos do critério de simplicidade. Em linhas gerais, mais simples é uma hipótese h_1, maior é a sua probabilidade inicial. Assumindo uma partição de hipóteses alternativas $\{h_1, h_2, h_3, \cdots, h_n\}$, quando todas elas correspondem igualmente bem aos outros três critérios — escopo, encaixe com a evidência de fundo e poder explanatório — h_1 terá probabilidade inicial maior do que as outras porque h_1 é a mais simples delas. Em acréscimo a essa estratégia, Swinburne ainda alega que h_1 terá probabilidade posterior maior do que a probabilidade posterior de cada uma das suas competidoras. Formulamos, todavia, duas razões contra a sua proposta. Primeira, como apontamos, normalmente hipóteses explanatórias cumprem os aspectos de simplicidade em níveis diferentes. Mas, afinal, há algum fator de simplicidade que prevalece sobre os outros? Swinburne não fornece nenhum argumento a favor de algum meta-critério que estabeleça uma hierarquia de importância entre os seis aspectos de simplicidade. Segunda, mesmo que h_1 tenha desempenho global *ótimo* em relação a todos os fatores de simplicidade, não é claro como eles aumentam a probabilidade inicial de h_1, $Pr(h_1 \mid k)$, e, consequentemente, a sua probabilidade posterior $Pr(h_1 \mid e \wedge k)$. Se o critério de simplicidade não é decidível para probabilidades iniciais, por que ele seria no que diz respeito às probabilidades posteriores? Além dessa crítica, sugerimos outros três problemas e objeções: a ausência de um argumento de Swinburne a favor da tese de coerência probabilística, a principal tese do probabilismo, a suposição da onisciência lógica e o problema da evidência antiga. Nenhuma delas é respondida de modo satisfatório por Swinburne e, advertidamente, as duas últimas continuam a ser críticas importantes a teorias Bayesianas em geral.

Resta-nos apontar algumas direções que consistem em prospectos de pesquisa a serem desenvolvidos na área de epistemologia formal. Sobre o problema da evidência antiga: novas medidas de confirmação ou reformulações de estratégias clássicas precisam ser empreendidas. Sobre simplicidade: a questão é determinar formalmente como os fatores podem ter impacto sobre as probabilidades iniciais e, eventualmente, sobre as probabilidades posteriores. Sobre argumentos em favor do probabilismo: os argumentos de James Joyce (1998 e 2009) e de Richard Pettigrew (2013, 2015 e 2016 (*forthcoming*)) constituem defesas promissoras em prol dessa tese. Sobre onisciência lógica: jogando com o título de um artigo de Richard Jeffrey (1983a), *why not a Bayesianism with a human face*?

A Apêndice

Equivalências lógicas

$(a)\ p \equiv [(p \wedge q) \vee (p \wedge \neg q)]$

$(b)\ q \equiv [(p \wedge q) \vee (\neg p \wedge q)]$

$(c)\ (p \vee q) \equiv [(p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge q)]$

Axiomas do Cálculo de Probabilidades

Para quaisquer p e q de $\mathbb{F}$ tal que $Pr : \mathbb{F} \longrightarrow \mathbb{R}$:

- (1) $Pr(p) \geq 0$;
- (2) Se p é uma tautologia, então $Pr(p) = 1$;
- (3) $Pr(p \vee q) = Pr(p) + Pr(q)$ se p e q são mutuamente exclusivas.

Definição de Probabilidade Condicional

$$Pr(p \mid q) = \frac{Pr(p \wedge q)}{Pr(q)},\ \textit{dado que } Pr(q) \neq 0$$

Teorema 1: $Pr(\neg p) = 1 - Pr(p)$.

Prova: Pelo axioma (3), $Pr(p \vee \neg p) = Pr(p) + Pr(\neg p)$, uma vez que p e $\neg p$ são mutuamente exclusivas. Pelo axioma (2), $Pr(p \vee \neg p) = 1$, pois $p \vee \neg p$ é uma tautologia. Assim, $Pr(p) + Pr(\neg p) = 1$. Portanto, $Pr(\neg p) = 1 - Pr(p)$. ■

Teorema 2: Se q é uma contradição lógica, então $Pr(q) = 0$.

Prova: Suposição de que q é uma contradição e p é uma tautologia. Assim sendo, $p \vee q$ é uma tautologia. Pelo axioma (2), $Pr(p \vee q) = 1$. Assim, pelo axioma (3), $Pr(p \vee q) = Pr(p) + Pr(q)$, pois p e q são mutuamente exclusivas. Por isso, $Pr(p) + Pr(q) = 1$. Pelo axioma (2), $Pr(p) = 1$. Portanto, $Pr(q) = 0$. ■

Teorema 3: Se $p \equiv q$, então $Pr(p) = Pr(q)$.

Prova: Suposição de que $p \equiv q$. Se $p \equiv q$, então $p \vee \neg q$ é uma tautologia. Pelo axioma

(2), $Pr(p \vee \neg q) = 1$. Pelo axioma (3), $Pr(p) + Pr(\neg q) = Pr(p \vee \neg q)$, 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 4: $Pr(p \vee q) = Pr(p) + Pr(q) - Pr(p \wedge q)$.

Prova: Pelo teorema 3 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 3 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 3 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, $Pr(p \vee q) = Pr(p \wedge q) + Pr(p \wedge \neg q) + Pr(\neg p \wedge q)$ (m). 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 (m), 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 5: Se $p \vDash q$, então $Pr(p) \leq Pr(q)$.

Prova: Suposição de que $p \vDash q$. Se $p \vDash q$, então $p \equiv (p \wedge q)$. Pelo teorema 3, $Pr(p) = Pr(p \wedge q)$ (n). Pela equivalência (b) e teorema 3, $Pr(q) = Pr[(p \wedge q) \vee (\neg p \wedge q)]$. 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 que $Pr(q) = Pr(p \wedge q) + Pr(\neg p \wedge q)$. 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 6: Para qualquer p, $Pr(p) \leq 1$.

Prova: Pelo axioma (1), $Pr(p) \geq 0$. Pelo teorema 5, se $p \vDash a$, então $Pr(p) \leq Pr(a)$. Para qualquer tautologia a, $p \vDash a$. Assim, $Pr(p) \leq Pr(a)$. Pelo axioma (2), $Pr(a) = 1$. Portanto, $Pr(p) \leq 1$. ■

Teorema 7: Se $p \vDash q$, então $Pr(q \mid p) = 1$.

Prova: Suposições de que $p \vDash q$ e $Pr(p) \neq 0$. Pela definição de probabilidade condicional, $Pr(q \mid p) = \frac{Pr(q \wedge p)}{Pr(p)}$. Se $p \vDash q$, então $p \equiv (p \wedge q)$. Pelo teorema 3, $Pr(p) = Pr(p \wedge q)$. Uma vez que $(p \wedge q) \equiv (q \wedge p)$, $Pr(q \mid p) = \frac{Pr(p)}{Pr(p)}$. Portanto, $Pr(q \mid p) = 1$. ■

Teorema 8: Para quaisquer p e q, $Pr(p) \geq Pr(p \wedge q)$.

Prova: Pelo teorema 3 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, $Pr(p) = Pr(p \wedge q) + Pr(p \wedge \neg q)$. Se $Pr(p \wedge \neg q) > 0$, então $Pr(p) > Pr(p \wedge q)$. Se $Pr(p \wedge \neg q) = 0$, então $Pr(p) = Pr(p \wedge q)$. Portanto, $Pr(p) \geq Pr(p \wedge q)$. ■

Teorema 9: $Pr(p) = [Pr(q) \times Pr(p \mid q)] + [Pr(\neg q) \times Pr(p \mid \neg q)]$, dado que $1 > Pr(q) > 0$.

Prova: Suposição de que $1 > Pr(q) > 0$. Pelo teorema 3 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, $Pr(p \mid q) = \frac{Pr(p \wedge q)}{Pr(q)}$. Assim, $Pr(p \wedge q) = Pr(p \mid q) \times Pr(q)$. $Pr(p \mid \neg q) = \frac{Pr(p \wedge \neg q)}{Pr(\neg q)}$ pela definição de probabilidade condicional. Assim, $Pr(p \wedge \neg q) = Pr(p \mid \neg q) \times Pr(\neg q)$. Logo, substituindo em (g), $Pr(p) = [Pr(q) \times Pr(p \mid q)] + [Pr(\neg q) \times Pr(p \mid \neg q)]$. ■

Teorema 10: $Pr(p \wedge q \mid r) = Pr(p \mid q \wedge r) \times Pr(q \mid r)$.

Prova: Assumindo que $Pr(q \wedge r) \neq 0$, segue-se pela definição de probabilidade condicional que $Pr(p \mid q \ \wedge \ r) = \frac{Pr(p \ \wedge \ q \ \wedge \ r)}{Pr(q \ \wedge \ r)}$. Por conseguinte, $Pr(p \wedge q \wedge r) = Pr(p \mid q \wedge r) \times Pr(q \wedge r)$ (j). Assumindo que $Pr(r) \neq 0$, $Pr(q \mid r) = \frac{Pr(q \ \wedge \ r)}{Pr(r)}$ pela definição de probabilidade condicional. Assim sendo, $Pr(q \wedge r) = Pr(q \mid r) \times Pr(r)$ (v). Pela definição de probabilidade condicional, segue-se que $Pr(p \wedge q \mid r) = \frac{Pr(p \ \wedge \ q \ \wedge \ r)}{Pr(r)}$. Usando (j) na última fórmula, $Pr(p \wedge q \mid r) = \frac{Pr(p \ \mid \ q \ \wedge \ r) \ \times \ Pr(q \ \wedge \ r)}{Pr(r)}$. Usando (v), $Pr(p \wedge q \mid r) = \frac{Pr(p \ \mid \ q \ \wedge \ r) \ \times \ Pr(q \ \mid \ r) \ \times \ Pr(r)}{Pr(r)}$. Portanto, conclui-se que $Pr(p \wedge q \mid r) = Pr(p \mid q \wedge r) \times Pr(q \mid r)$. ■

Teorema 11: Se $Pr(q) = 1$, então $Pr(p \wedge q) = Pr(p)$.

Prova: Suposição de que $Pr(q) = 1$. Pelo teorema 9, segue-se que $Pr(p) = [Pr(q) \times Pr(p \mid q)] + [Pr(\neg q) \times Pr(p \mid \neg q)]$. Assim, $Pr(p) = [1 \times Pr(p \mid q)] + [0 \times Pr(p \mid \neg q)]$, uma vez que $Pr(\neg q) = 1 - Pr(q)$. Assim, $Pr(p) = Pr(p \mid q)$. Pela definição de probabilidade condicional, $Pr(p \mid q) = \frac{Pr(p \ \wedge \ q)}{Pr(q)}$. Segue-se que $Pr(p \mid q) = Pr(p \wedge q)$, pois $Pr(q) = 1$. Portanto, $Pr(p) = Pr(p \wedge q)$. ■

Teorema 12:

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

dado que $Pr(e) > 0$.

Prova: Suposição de que $Pr(e) > 0$. Pela definição de probabilidade condicional, $Pr(h \mid e) = \frac{Pr(h \ \wedge \ e)}{Pr(e)}$ e $Pr(e \mid h) = \frac{Pr(e \ \wedge \ h)}{Pr(h)}$. 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 9, $Pr(e) = [Pr(h) \times Pr(e \mid h)] + [Pr(\neg h) \times Pr(e \mid \neg h)]$. 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 13:

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

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

Prova: Suposição de que $Pr(e \mid k) > 0$. Pela definição de probabilidade condicional,

$Pr(h \mid e \wedge k) = \frac{Pr(h \wedge e \wedge k)}{Pr(e \wedge k)}$ e $Pr(e \mid h \wedge k) = \frac{Pr(e \wedge h \wedge k)}{Pr(h \wedge k)}$. 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, $Pr(h \mid e \wedge k) = \frac{Pr(e \mid h \wedge k) \times Pr(h \wedge k)}{Pr(e \wedge k)}$. Pela definição de probabilidade condicional, $Pr(h \mid k) = \frac{Pr(h \wedge k)}{Pr(k)}$ e $Pr(e \mid k) = \frac{Pr(e \wedge k)}{Pr(k)}$. Assim, segue-se que $Pr(h \wedge k) = Pr(h \mid k) \times Pr(k)$ 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

ALSTON, W. **Epistemic Justification: Essays in the Theory of Knowledge**. Ithaca and London: Cornell University Press, 1989.

ALSTON, W. Swinburne on faith and belief. In: PADGETT, A. G. (Ed.). **Reason and the Christian Religion**. Oxford: Clarendon Press, 1994. p. 21–37.

ALSTON, W. **Beyond "Justification": Dimensions of Epistemic Evalution**. Ithaca and London: Cornell University Press, 2005.

BAKER, A. Simplicity. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 28 de março de 2016: http://plato.stanford.edu/entries/simplicity/, 2010.

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.

BERTRAND, J. **Calcul Des Probabilités**. Toronto: University of Toronto Libraries, 2011.

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.

CHISHOLM, R. **Theory of Knowledge**. 2nd ed. Englewood Cliffs: Prentice Hall, 1977.

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.

COHEN, J. L. **An Essay on Belief and Acceptance**. Oxford: Oxford University Press, 1992.

DEMPSTER, A. P. A generalization of bayesian inference. **Journal of the Royal Statistical Society**, (Series B, Methodological), v. 30, p. 205–247, 1968.

DUBOIS, D.; PRADE, H. **Possibility Theory**. New York and London: Plenum Press, 1988.

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.

ELGA, A. Subjective probabilities should be sharp. **Philosophers' Imprint**, v. 10, n. 5, p. 1–11, 2010.

FELDMAN, R.; CONEE, E. Internalism defended. **American Philosophical Quarterly**, v. 38, n. 1, p. 1–18, 2001.

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.

FITELSON, B.; HARTMANN, S. A new garber-style solution to the problem of old evidence. **Philosophy of Science**, n. 4, p. 712–717, 2015.

FOLEY, R. Beliefs, degrees of belief, and the lockean thesis. In: HUBER, F.; SCHMIDT-PETRI, C. (Ed.). **Synthese Library 342, Degrees of Belief**. Dordrecht: Synthese, 2009. p. 37–47.

GARBER, D. Old evidence and logical omniscience in bayesian confirmation theory. In: EARMAN, J. (Ed.). **Testing Scientific Theories. Midwest Studies in the Philosophy of Science, Vol. X**. Minneapolis: University of Minnesota Press, 1983. p. 99–131.

GETTIER, E. L. Is justified true belief knowledge? **Analysis**, v. 23, n. 6, p. 121–123, 1963.

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

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

GOLDMAN, A. What is justified true belief? In: PAPPAS, G. S. (Ed.). **Justification and Knowledge**. Dordrecht: D. Reidel, 1979. p. 1–23.

GOLDMAN, A. **Epistemology and Cognition**. Cambridge, MA: Harvard University Press, 1986.

GOODMAN, N. **Fact, Fiction, and Forecast**. Cambridge, MA: Harvard University Press, 1983.

HACKING, I. **The Logic of Statistical Inference**. Cambridge, UK: Cambridge University Press, 1965.

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

HÁJEK, A. What conditional probability could not be? **Synthese**, v. 137, n. 3, p. 273–323, 2003.

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**. Último Acesso em: 12 de Janeiro de 2016: http://plato.stanford.edu/entries/probability-interpret/, 2011.

HÁJEK, A. Pascal's wager. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 10 de Janeiro de 2016: http://plato.stanford.edu/entries/pascal-wager/, 2012.

HANDFIELD, T. **A Philosophical Guide to Chance: Physical Probability**. Cambridge: Cambridge University Press, 2012.

HAWTHORNE, J. The lockean thesis and the logic of belief. In: HUBER, F.; SCHMIDT-PETRI, C. (Ed.). **Synthese Library 342, Degrees of Belief**. Dordrecht: Synthese, 2009. p. 49–74.

HEMPEL, C. G. **Aspects of Scientific Explanation and Other Essays in the Philosophy of Science**. New York: The Free Press, 1965.

HENDRICKS, V. **Mainstream and Formal Epistemology**. Cambridge: Cambridge University Press, 2006.

HENDRICKS, V.; SYMONS, J. Epistemic logic. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 20 de Janeiro de 2016: http://plato.stanford.edu/entries/logic-epistemic/, 2006.

HINTIKKA, J. **Knowledge and Belief: An Introduction to the Logic of the Two Notions**. Ithaca and London: Cornell University Press, 1962.

HORWICH, P. **Probability and Evidence**. Cambridge, UK: Cambridge University Press, 1982.

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.). **Degrees of Belief, Synthese Library 342**. Dordrecht: Synthese, 2009. p. 1–33.

HUBER, F. Confirmation and induction. In: FIESER, J.; DOWDEN, B. (Ed.). **Internet Encyclopedia of Philosophy**. Último Acesso em: 17 de janeiro de 2016: http://www.iep.utm.edu/conf-ind/, 2010.

HUBER, F.; SCHMIDT-PETRI, C. **Degrees of Belief, Synthese Library 342**. Dordrecht: Synthese, 2009.

JAYNES, E. T. **Probability Theory: the logic of science**. Cambridge, UK: Cambridge University Press, 2003.

JEFFREY, R. Bayesianism with a human face. In: EARMAN, J. (Ed.). **Testing Scientific Theories. Midwest Studies in the Philosophy of Science, Vol. X**. Minneapolis: University of Minnesota Press, 1983a. p. 133–156.

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

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.

JOYCE, J. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In: HUBER, F.; SCHMIDT-PETRI, C. (Ed.). **Synthese Library 342, Degrees of Belief**. Dordrecht: Synthese, 2009. p. 263–297.

JOYCE, J. The development of subjective bayesianism. In: GABBAY, D.; HARTMANN, S.; WOODS, J. (Ed.). **Handbook of the History of Logic, vol. 10, Inductive Logic**. Amsterdan: North-Holland, 2011. p. 415–475.

KAPLAN, M. **Decision Theory as Philosophy**. Cambridge: Cambridge University Press, 1996.

KELLY, T. Epistemic rationality as instrumental rationality: A critique. **Philosophy and Phenomenological Research**, v. 66, n. 3, p. 612–640, 2003.

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

KEYNES, J. M. **A Treatise on Probability**. London: Macmillan, 1921.

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

LAPLACE, P. S. **A Philosophical Essay on Probabilities**. New York: Dover, 1951.

LEWIS, D. **Counterfactuals**. Malden, MA: Blackwell Publishing, 1973.

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.

LEWIS, D. Humean supervenience debugged. **Mind**, v. 103, n. 412, p. 473–490, 1994.

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

MAHER, P. Depragmatized dutch book arguments. **Philosophy of Science**, v. 64, n. 2, p. 291–305, 1997.

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

MISES, R. v. **Probability, Statistics and Truth**. New York: Macmillan, 1957.

NEIVA, A. Probabilismo e bayesianismo em epistemologia. **PERI**, v. 07, n. 02, p. 45–69, 2015.

PETTIGREW, R. Epistemic utility and norms for credences. **Philosophy Compass**, v. 8, n. 10, p. 897–908, 2013.

PETTIGREW, R. Epistemic utility arguments for probabilism. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 26 de fevereiro de 2016: http://plato.stanford.edu/entries/epistemic-utility/, 2015.

PETTIGREW, R. **Accuracy and the Laws of Credence**. Oxford: Oxford University Press, 2016 (*forthcoming*).

POPPER, K. The propensity interpretation of probability. **The Britisth Journal for the Philosophy of Science**, v. 10, n. 37, p. 25–42, 1959.

POPPER, K. **The Logic of Scientific Discovery**. London and New York: Routledge, 2002.

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.

REICHENBACH, H. **The Theory of Probability**. Berkeley, CA: University of California Press, 1949.

RÉNYI, A. **Foundations of Probability**. San Francisco: Holden-Day Inc, 1970.

ROSA, L. A model-theoretic semantics for attributions of offline-rationality. **manuscript**, Disponível em http://www.luisrosa.org, *forthcoming*.

SAVAGE, L. **The Foundations of Statistics**. New York: Wiley Publications in Statistics, 1954.

SCHWITZGEBEL, E. Belief. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 18 de janeiro de 2016: http://plato.stanford.edu/entries/belief/, 2015.

SHAFER, G. **A Mathematical Theory of Evidence**. Princeton, NJ: Princeton University Press, 1976.

SHOGENJI, T. The degree of epistemic justification and the conjunction fallacy. **Synthese**, v. 184, n. 1, p. 29–48, 2012.

SOBER, E. **Ockham's Razors: A User's Manual**. Cambridge UK: Cambridge University Press, 2015.

STEELE, K.; STEFÁNSSON, H. O. Decision theory. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 20 de fevereiro de 2016: http://plato.stanford.edu/entries/decision-theory/, 2015.

STEUP, M. Epistemology. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 15 de janeiro de 2016: http://plato.stanford.edu/entries/epistemology/, 2005.

STREVENS, M. Objective probability as a guide to the world. **Philosophical Studies**, v. 95, n. 3, p. 243–275, 1999.

STREVENS, M. **Notes on Bayesian Confirmation Theory**. Disponível em: http://www.nyu.edu/classes/strevens/BCT/BCT.pdf, *manuscript*.

SWINBURNE, R. **Faith and Reason**. 1st ed. Oxford: Clarendon Press, 1981.

SWINBURNE, R. **Simplicity as Evidence of Truth**. Milwaukee: Marquette University Press, 1997.

SWINBURNE, R. **Epistemic Justification**. Oxford: Clarendon Press, 2001.

SWINBURNE, R. Introduction. In: SWINBURNE, R. (Ed.). **Bayes's Theorem**. Oxford: Oxford University Press. Proceedings of the British Academy 113, 2002. p. 01–20.

SWINBURNE, R. **The Existence of God**. 2nd ed. Oxford: Clarendon Press, 2004.

SWINBURNE, R. **Faith and Reason**. 2nd ed. Oxford: Clarendon Press, 2005.

SWINBURNE, R. **Mind, Brain, and Free Will**. Oxford: Oxford University Press, 2013.

TALBOTT, W. Bayesian epistemology. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 19 de janeiro de 2016: http://plato.stanford.edu/entries/epistemology-bayesian/, 2008.

TITELBAUM, M. G. **Quitting Certainties: A Bayesian Framework Modeling Degrees of Belief**. Oxford: Oxford University Press, 2013.

VINEBERG, S. Dutch book arguments. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 19 de janeiro de 2016: http://plato.stanford.edu/entries/dutch-book/, 2011.

WEISBERG, J. Varieties of bayesianism. In: GABBAY, D.; HARTMANN, S.; WOODS, J. (Ed.). **Handbook of the History of Logic, vol. 10, Inductive Logic**. Amsterdan: North-Holland, 2011. p. 477–551.

WEISBERG, J. Formal epistemology. In: ZALTA, E. (Ed.). **The Stanford Encyclopedia of Philosophy**. Último Acesso em: 23 de Janeiro de 2016: http://plato.stanford.edu/entries/formal-epistemology/, 2015.

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

WILLIAMSON, T. **Knowledge and its Limits**. Oxford: Oxford University Press, 2000.

ZADEH, L. A. Fuzzy sets as a basis for a theory of possibility. **Fuzzy Sets and Systems**, Vol. 1, p. p. 3–28, 1978.