Ricardo Gentil de Araújo Pereira

Interpretação e análise do problema da indução sob uma visão fundamentada em teoria de conjuntos e teoria de probabilidades

Natal
2012

Ricardo Gentil de Araújo Pereira

Interpretação e análise do problema da indução sob uma visão fundamentada em teoria de conjuntos e teoria de probabilidades

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal do Rio Grande do Norte, para a obtenção de título de Mestre em Filosofia.

Linha de pesquisa: Lógica e Filosofia Formal

Orientador: Daniel Durante Pereira Alves

Natal
2012

Catalogação da Publicação na Fonte.
Universidade Federal do Rio Grande do Norte.
Biblioteca Setorial do Centro de Ciências Humanas, Letras e Artes (CCHLA).

Pereira, Ricardo Gentil de Araújo.
Interpretação e análise do problema da indução sob uma visão fundamentada em teoria de conjuntos e teoria de probabilidades / Ricardo Gentil de Araújo Pereira. – 2012.
108 f. :il. -

Dissertação (Mestrado em Filosofia) – Universidade Federal do Rio Grande do Norte. Centro de Ciências Humanas, Letras e Artes. Programa de Pós-Graduação em Filosofia, Natal, 2012.
Orientador: Prof. Dr. Daniel Durante Pereira Alves.
Área de concentração: Lógica e Filosofia formal.

1. Lógica indutiva. 2. Problema da indução. 3. Probabilidade. I. Alves, Daniel Durante Pereira. II. Universidade Federal do Rio Grande do Norte. III. Título.

RN/BSE-CCHLA CDU 164

Comissão Julgadora:

Prof. Dr.
José Eduardo de Almeida Moura

Prof. Dr.
Giovanni da Silva de Queiroz

Prof. Dr.
Daniel Durante Pereira Alves

Agradecimentos

Esse trabalho de pesquisa só foi possível graças a toda uma cadeia de suporte da qual pude usufruir e a que sou muito grato. Dentre as inúmeras ajudas, tanto em termos materiais quanto emocionais, agradeço explicitamente aos alunos da base de lógica, a Thiare e demais membros da coordenação da pós-graduação, à Cláudia (secretaria da graduação), à professora Maria da Paz e à ajuda fundamental do meu orientador Daniel Durante e dos professores José Eduardo e Bruno Vaz.

Por último, mas não menos importante, agradeço especialmente aos meus pais, ao grande amigo Allan Medeiros pela inoxidável ajuda e incentivo, e à Dária Baísa por ter me acompanhado e motivado durante todo o processo, além de revisado incansavelmente todo o texto da dissertação.

Resumo

O seguinte trabalho consiste na interpretação e análise do problema da indução sob uma visão fundamentada em teoria de conjuntos e teoria de probabilidades como base para a solução de suas implicações filosóficas negativas relativas aos sistemas de lógica indutiva de maneira geral. Devido à importância do problema e aos desenvolvimentos recentes nos referidos campos de conhecimento (início do século 20), bem como às relações visíveis entre eles e o processo de inferência indutivo, tem-se aberto um campo de possibilidades relativamente inexplorado e promissor. O ponto-chave para o estudo consiste na modelagem do processo de aquisição de informação usando conceitos de teoria de conjuntos, seguido por um tratamento usando teoria de probabilidades. Ao longo do estudo foi possível identificar, como obstáculos principais à justificação probabilística, tanto o problema da definição do conceito de probabilidade quanto do de racionalidade, além da sutil conexão entre ambos. Essa constatação permitiu um maior cuidado na escolha do critério de racionalidade a ser considerado no intuito de viabilizar o tratamento do problema por meio de situações-exemplo específicas, mas sem a perda de suas características originais, de modo que as conclusões obtidas possam ser estendidas a casos clássicos como o relativo à dúvida sobre a continuidade do nascer do sol.

Palavras-chave: problema da indução, lógica indutiva, probabilidade

Abstract

The following work is to interpret and analyze the problem of induction under a vision founded on set theory and probability theory as a basis for solution of its negative philosophical implications related to the systems of inductive logic in general. Due to the importance of the problem and the relatively recent developments in these fields of knowledge (early 20th century), as well as the visible relations between them and the process of inductive inference, it has been opened a field of relatively unexplored and promising possibilities. The key point of the study consists in modeling the information acquisition process using concepts of set theory, followed by a treatment using probability theory. Throughout the study it was identified as a major obstacle to the probabilistic justification, both: the problem of defining the concept of probability and that of rationality, as well as the subtle connection between the two. This finding called for a greater care in choosing the criterion of rationality to be considered in order to facilitate the treatment of the problem through such specific situations, but without losing their original characteristics so that the conclusions can be extended to classic cases such as the question about the continuity of the sunrise.

Keywords: problem of induction, inductive logic, probability

Lista de Tabelas

1.1 Características das amostras representativas do problema 38

3.1 Sequência pior caso 71

3.2 Tabela de evolução de escolhas confome hábito 78

3.3 Possibilidades de fabricação da moeda, conforme informações fornecidas 84

3.4 Probabilidades iniciais das hipóteses 87

3.5 Atualização das probabilidades das hipóteses no caso 1 88

3.6 Atualização das probabilidades das hipóteses no caso 2 89

3.7 Inferências relativas à sequência-exemplo 92

3.8 Inferências relativas ao pior caso 93

Sumário

1 Introdução **17**

1.1 A indução 21

1.2 Histórico 22

1.3 O Problema 24

1.3.1 Explicação 24

1.3.2 *Explicandum* 24

1.3.3 *Explicatum* 38

2 As implicações e o modelo **43**

2.1 O desmoronamento do conhecimento científico 43

2.2 A impossibilidade do "conhecimento probabilístico" 46

2.3 A Modelagem e as Ferramentas 47

2.3.1 O exemplo prático: lançamento de moeda 47

2.3.2 Modelagem 54

2.3.3 Teoria de Conjuntos 57

2.3.4 Teoria de Probabilidades 60

3 Análise e considerações finais **67**

3.1 A dinâmica do agente 67

3.2 Identificando a instância do problema no âmbito do modelo 72

3.3 O tratamento do problema 73

3.3.1 Definindo racionalidade 75

3.3.2 A racionalidade da Teoria de Probabilidades 80

3.3.3 Análise quantitativa 83
3.4 Considerações Finais 94
3.4.1 Sobre o nascer do sol 96
3.4.2 Sobre o hábito 98

Referências Bibliográficas **101**

Appendices **107**

A Programa coin-all.py **107**

Capítulo 1

Introdução

O presente trabalho representa uma abordagem do clássico problema da indução visto no âmbito do processo de aquisição de conhecimento. Para tanto, foi considerado o processo de aprendizado individual como ponto de partida, seguido de um breve estudo histórico e da identificação mais precisa do objeto a ser tratado.

Mesmo nessa etapa inicial foram usados ferramentas e conceitos atuais que se julga poderem facilitar sua compreensão e tratamento, mas cuja possibilidade de interpretações divergentes torna necessária uma explicação antecipada, de modo que o leitor possa entender em que sentido estão sendo usados ao longo do texto, ainda que em um momento apropriado sejam permitidas modificações ou adaptações devidamente explicitadas.

O termo conhecimento, por exemplo, não está sendo tomado no sentido clássico de crença verdadeira justificada (ou ainda reforçada por algum critério visando evitar o problema de Gettier (Gettier, 1963)), mas apenas como aquisição de informação em conformidade com os critérios disponíveis para o agente cognoscível considerado, ainda que esses critérios representem apenas de maneira imperfeita a busca pelo estabelecimento da crença verdadeira[1], bem como sua manutenção.

Assim, ao abrir mão da eliminação da possibilidade do erro, que é a principal característica apontada por Platão e que daria um valor instrumental à modalidade

[1]Essa idéia é similar à caracterização do conhecimento como descrevendo o sucesso no empreendimento de busca da verdade, conforme exposto em (BonJour, 2009, pg. 30)

de crenças que estariam "atadas à verdade" (Pritchard and Turri, 2012), poder-se-ia optar alternativamente pelo uso de *crença justificada*, de forma a apenas indicar a adequação aos critérios do agente. Ao fazer isso, entretanto, perderíamos o contato com o ideal de certeza presente na ideia filosófica de conhecimento e, ainda que em diferentes graus, também em seu emprego científico e popular.

Importante observar que o conhecimento abordado é do tipo proposicional (Steup, 2012), principalmente em relação à predição de algum evento, como por exemplo: S "sabe" que P, onde P é o próximo resultado do lançamento de determinada moeda. As aspas usadas na declaração da sentença são apenas para chamar a atenção para o fato de que a forma simples e explícita de tal exemplo probabilístico dá uma falsa impressão de que há uma patente diferença entre ele e as situações filosóficas, científicas e populares mencionadas.

Isso não é verdade, entretanto, já que é possível contemplar casos com nível de confiança variando de 0 (nenhuma) a 1 (certeza), englobando qualquer valor de probabilidade que pensemos poder servir de limite a partir do qual uma crença passaria a ser considerada virtualmente certa e receberia o nobre rótulo de conhecimento. Isso posto, também faz-se prudente um esclarecimento a respeito dos seguintes termos:

- Lógica
- Indução
- Aquisição de conhecimento
- Predição
- Probabilidade
- Evolução

Ao usarmos o termo lógica, estamos nos referindo a um sistema formal de inferências por meio de regras de manipulação simbólica, usado como ferramenta para ajudar no raciocínio a respeito de determinados problemas.

Embora tradicionalmente a **indução** (ou raciocínio indutivo) seja entendida como a inferência de uma afirmação universal a partir de instâncias particulares, nesse estudo seguiremos uma visão aproximada à de Rudolf Carnap, David Stove e Colin Howson, conforme (Carnap, 1962), (Stove, 1986) e (Howson, 2003), respectivamente, compreendendo-a como o processo de inferência de um evento contingente a partir de um conjunto já conhecido. Essa perspectiva é compatível com a visão de David Hume sobre inferência de eventos futuros a partir de observações do passado (Hume, 1978), além de preservar o núcleo do problema, conforme explicado em 1.3.3.

Nesse contexto, a **aquisição de conhecimento** será entendida simplesmente como a adição de elementos ao conjunto de informações, assim como o "reconhecimento" de características (tanto do conjunto, quanto dos elementos) relevantes para a situação descrita.

A partir dessas considerações iniciais sobre lógica e indução, portanto, uma lógica indutiva consiste em um sistema formal, construído segundo regras de manipulação simbólica que, diferentemente da lógica clássica, em vez de preservar a validade, preserva uma racionalidade simplificada. [2]

Apesar das várias críticas a essa ideia, as principais objeções dizem respeito à valoração de algumas das premissas necessárias ao argumento indutivo probabilístico: as probalilidades iniciais. Se considerarmos, por exemplo, que no caso de uma inferência dedutiva via *modus ponens* do tipo que leva de A e $A \to B$ para B, a incerteza quanto ao valor de verdade de uma das premissas não afeta a validade formal do argumento, mas apenas a sua correção, o mesmo se aplicaria no caso da incapacidade de estabelecer as probabilidades iniciais de uma inferência indutiva. Por último, mesmo considerando as objeções quanto a aplicabilidade prática, há algumas defesas combatendo a ideia de que a probabilidade seria incapaz de ajudar na resolução do problema ainda segundo a argumentação de Hume (Stove, 1965), bem como abordando as críticas mais atuais como as referidas questões sobre a atribuição de probabilidades iniciais (Franklin,

[2] Essa "racionalidade simplificada", considerada como um subconjunto legítimo dos critérios necessários, ainda que não suficientes, para a caracterização de uma agente como racional, é representada pelos desiderata de Pólia-Cox, conforme 2.3.4.

2001).

Quanto a **predição**, embora uma visão em termos de teoria de conjuntos seja capaz de abstrair o conceito temporal normalmente atrelado ao termo, será entendida como o processo de estipular antecipadamente o resultado de um experimento, como o lançamento de uma moeda ou o sorteio de bolas de uma urna, que serão os casos considerados como referências.

Esses casos escolhidos remetem claramente ao uso da teoria de probabilidades em seu tratamento, mas como o próprio termo **probabilidade** ainda é controverso, faz-se necessário explicitar em que sentido está sendo usado nesse trabalho.

Diante das várias possibilidades de interpretação (Hájek, 2012) e da aplicação restrita aos conjuntos finitos a serem considerados, serão favorecidas tanto a visão clássica que a define numericamente como razão entre casos favoráveis e casos possíveis (relativa a dado experimento), quanto a visão de probabilidade como representando hipótese sobre essa razão (calculada a partir da informação disponível), também usada para representar o grau de crença na ocorrência dos casos favoráveis mencionados.

No clássico problema de *Monty Hall*[3], por exemplo, enquanto a probabilidade de ganhar o prêmio para a estratégia de sempre trocar de porta é $2/3$, a crença da maioria das pessoas é a de que a troca é indiferente, o que pode ser representado como atribuição de probabilidade de $1/2$ para a chance de ganhar mediante troca.

Ainda pensando sobre o problema de *Monty Hall* é possível ver um ponto importante relacionado à opção por considerar conjuntos finitos nas inferências, bastando que imaginemos o caso de infinitas portas. Ora: sendo o conjunto infinito, todas receberiam probabilidade zero segundo a regra básica mencionada anteriormente. Por isso é importante que se verifique a preservação das características principais do problema da indução mesmo diante de tal opção, como será visto mais adiante.

Se a probabilidade representa informação sobre o sistema considerado, o problema

[3]Suponha que você esteja em um programa de jogos e lhe seja dada a opção de escolher entre três portas: atrás de uma porta tem um carro; atrás das outras, cabras. Digamos que você escolha a porta Nº 1 (mas a porta não é aberta) e o anfitrião, que sabe o que está por trás das portas, abre a porta Nº 3, que tem uma cabra. Ele então lhe diz: “Você quer trocar para a porta Nº 2?”. É vantagem mudar a sua escolha?(Savant, 1990; Selvin, 1975)

da indução também pode ser entendido como a impossibilidade de distinção entre decisões diferentes tomadas com base nessa informação. Ao fazer essa comparação entre decisões (ou regras) e sua influência sobre determinado agente, pode-se pensar em **evolução** – num sentido genérico – como a mudança sequencial nas "populações" de regras, assim como no caso biológico é entendida como o processo de mudança das propriedades dos grupos de organismos ao longo das gerações (Futuyma, 2005, pg. 2).

1.1 A indução

Do ponto de vista individual, tomamos decisões baseadas em raciocínio indutivo desde que nascemos e continuamos a fazê-lo durante toda a nossa vida, sem sequer darmos nome a esse processo ou questionarmos sua validade e alcance.

As inferências indutivas são o primeiro mecanismo de aquisição de conhecimento de que dispomos, inclusive sendo responsáveis pela homologação de "fontes confiáveis" de conhecimento indireto, como nossos pais. Entenda-se por homologação (reconhecimento ou aprovação) simplesmente a confirmação de determinada informação recebida, como a associação entre um determinado ato e as consequências boas ou más relacionadas à sua execução.[4] Apesar de a confiabilidade das informações ser bastante variável, ela tende a ser mais elevada à medida que as consequências do erro ou acerto tornam-se mais relevantes à nossa vida.

Decisões como não por a mão em uma lagarta-de-fogo, mesmo que nunca se tenha tido contato com tal animal, são tomadas rapidamente e, dependendo da gravidade do perigo, basta uma única experiência para aprendermos que não se deve repetir esse ato. Esse mecanismo instintivo é compreensivelmente mais forte conforme a relevância das implicações vitais da decisão a ser tomada frente a um fato novo, já que, por mais simples e impreciso que possamos concebê-lo, há um limite mínimo que deve ser obedecido: o da sobrevivência.

Pode-se fazer um paralelo interessante entre a história da aquisição de conheci-

[4] Por exemplo: a confirmação por meio da verificação dolorosa de que nossos pais estão corretos ao nos dizerem para não colocar a mão no fogo ou na tomada sob pena de nos machucar.

mento individual e a coletiva, apesar de apenas essa última ter permitido um tratamento mais aprofundado, acumulando evidências de sucesso e, ao mesmo tempo, questionamentos sobre uma justificação que se faz mais necessária proporcionalmente à formalização dos métodos cognitivos. Foi sob essa tensão entre resultados e carência de justificativa racional que o método indutivo gerou o problema da indução.

Como está sendo observado por meio de seu impacto no processo de aquisição de conhecimento, a abordagem foi estabelecida como sendo a análise de um experimento demonstrando aprendizagem, seguindo-se da identificação da ocorrência do problema e das características que o tornam inerente ao aprendizado em situações de incerteza.

Uma vez feito isso, partiu-se para a simplificação desse processo de forma a eliminar qualquer informação não relativa à constituição básica necessária para caracterização tanto do processo de aprendizagem quanto do problema da indução. Após a obtenção de um modelo simplificado, passou-se ao uso das ferramentas da Teoria de Conjuntos e da Teoria de Probabilidades, necessárias para seu tratamento e, finalmente, à análise da resolução do problema para o caso específico e suas implicações para o caso geral.

Por fim, apresenta-se a seguir uma breve abordagem histórica do problema da indução enquanto relacionado à tomada de decisão em situações de incerteza.

1.2 Histórico

Desde que o homem iniciou a organização do conhecimento e a modelagem de algo que seria chamado futuramente de método científico, tornou-se visível a aglutinação de filósofos e cientistas em torno de dois modos de aprendizagem. Esses métodos, chamados genericamente de indução e dedução, também podem ser entendidos como padrões de inferência e critérios de justificação epistêmica, sendo usados em etapas distintas do processo cognitivo e sofrendo críticas diferentes quanto aos seus alcances e adequações.

A indução é amplamente usada para predizer um evento particular baseado em ocorrências anteriores, assim como para compor um modelo geral a partir de determinada amostra. Apesar disso, várias questões foram levantadas desde os tempos do filósofo Sextus Empiricus, cujo texto em *Against the Logicians*, segundo a filósofa Ruth Weintraub, pode ser visto como análogo às questões formuladas por Hume em *An Enquiry concerning Human Understanding.*(Weintraub, 1995) Embora possa não ter sido o primeiro, certamente foi Hume que chamou a atenção para o princípio da indução por meio da derivação do "Problema da Indução".

O processo de inferência em situações de incerteza também é estudado por matemáticos há mais de 350 anos, iniciando com Gerolamo Cardano, com destaque para Blaise Pascal, Pierre de Fermat e o responsável pelo que é hoje considerado como a interpretação clássica de probabilidades: Pierre-Simon Laplace (Hájek, 2012). Esse estudo acabou por despertar o interesse em suas ligações com a lógica, inicialmente com os trabalhos de George Boole em seu livro *Laws of Thought* (1854), e de John Venn, em seu *The Logic of Chance* (1866).

Dentre as várias abordagens mais recentes do método indutivo, destaca-se a realizada pelo filósofo austríaco Karl Popper, em seu livro *The Logic of Scientific Discovery*, onde é questionada a necessidade científica da indução e, consequentemente, a sua caracterização como problema. Além de Popper, filósofos e estatísticos como Rudolf Carnap, John Maynard Keynes, Hans Reichenbach e Bruno de Finetti, abriram novas possibilidades de compreensão do processo de aquisição de conhecimento, como um todo, bem como do problema da indução em particular, resultando no desenvolvimento de sistemas de lógica indutiva dos quais destaca-se a Probabilidade Bayesiana.

Diante desse novo panorama, caracterizado por avanços na lógica e na matemática, esta pesquisa examinou duas abordagens específicas do problema da indução, representadas pelos livros *The Racionality of Induction*, de David Stove, e *Hume's Problem: Induction and the Justification of Belief*, de Colin Howson. Além desses dois livros, filosoficamente orientados, foi usado como referência o livro *Probability Theory: The Logic of Science*, do físico E.T. Jaynes, principalmente no que diz respeito aos fundamentos sobre os quais a probabilidade bayesiana se apóia, representados pelos dois primeiros capítulos, alçados a princípios fundamentais de racionalidade.

Além dos últimos livros mencionados, que constituem o tripé fundamental desse estudo, foi usado como referência importante o livro *Logical Foundations of Probability*, de Rudolf Carnap, inclusive a parte sobre *explicandum* e *explicatum*, que reflete o problema inicial do próprio conceito de probabilidade.

1.3 O Problema

O termo **problema da indução**, assim como ocorre com **probabilidade**, pode ser usado para nomear um conjunto de ideias relacionadas, passíveis de serem analisadas de maneiras distintas sem que se preocupe com uma definição exata, desde que as conclusões obtidas nos casos específicos não sejam extrapoladas sem as considerações pertinentes.

Supondo que há uma componente genérica do problema, responsável pela possibilidade de agrupamento das ideias em torno de um único termo, pode-se tentar a clarificação do conceito à maneira proposta em (Carnap, 1962), a partir de exemplos, exceções e do processo explicativo que consiste na transformação do conceito vago no conceito exato a ser tratado.

1.3.1 Explicação

Por explicação, entenda-se apenas a transformação do conceito vago ou, antes, do conjunto de ideias relacionadas, numa unidade exata e passível de utilização prática, ainda que se tenha de optar pela perda de correspondência rigorosa ao conjunto inicial (heterogêneo). Para facilitar referências, optou-se por manter a nomenclatura usada por Carnap para a ideia em seu estado bruto – *Explicandum* – e para o resultado após o processo de clarificação – *Explicatum*.

1.3.2 *Explicandum*

Como ponto de partida para a compreensão do que se está falando, o *explicandum* deve consistir pelo menos num conjunto de ideias que possam ser consideradas representantes do conceito abstrato, bem como de algumas exceções e explicações sobre a classificação dessas ideias como pertinentes ou não à amostra conceitual.

Desse modo, segue-se o exame das visões das principais personalidades que se ocuparam do tema, consideradas amostras representativas. Ao longo da exposição, serão identificadas características que possam ser consideradas componentes básicas do problema, caso estejam presentes em todas (ou na maioria) das amostras.

Amostragem

- Sextus Empiricus – É possível identificar pelo menos dois trechos nos quais é clara a problemática do método indutivo, tanto no seu aspecto amostral quanto à questão da circularidade, conforme apontado por Ruth Weintraub (Weintraub, 1995). São eles:

  > Quando eles propõe estabelecer o universal a partir dos particulares por meio da indução, o farão ou por meio da observação de todos ou de apenas alguns dos particulares. Mas se observam apenas alguns, a indução será insegura, já que os omitidos podem contrariar o universal; enquanto que se pretendem observar a todos, estarão almejando o impossível, já que os particulares são infinitos e indefinidos.[5] (Empiricus, 1933, tradução nossa).

  Nesse primeiro trecho, o problema é explicitamente identificado com a questão de se tentar chegar a conclusões universais a partir de indivíduos, característica que doravante chamaremos de *Particularidade* $\times$ *generalidade*.

  Ao mesmo tempo, é observado que a possibilidade de um indivíduo ainda desconhecido contrariar o universal, aliada à impossibilidade de conhecer a todos (no contexto relativo à cotação), implica na manutenção do caráter contingente quanto à verdade ou falsidade da aplicabilidade geral da característica. Ou seja: pode ser que ela seja, de fato, aplicável a todo o conjunto, mas também pode ser

[5]When they propose to establish the universal from the particulars by means of induction, they will effect this by a review of either all or some of the particulars. But if they review some, the induction will be insecure, since some of the particulars omitted in the induction may contravene the universal; while if they are to review all, they will be toiling at the impossible, since the particulars are infinite and indefinite.

que haja algum elemento que não a possua. Essa problemática será representada através da oposição *Contingência* × *necessidade*.

Ainda segundo Weintraub (Weintraub, 1995), Sextus questiona o uso de critérios de decisão entendidos de maneira geral, resultando em observações que podem ser aplicadas tanto à dedução quanto à indução. Nessa abordagem ele observa que um critério não poderia ser usado para justificar ele mesmo, o que resultaria em *caráter circular* (a estratégia usada por Hume) e, portanto, necessitaria de um critério anterior, gerando um regresso infinito.

- David Hume – Não tratou o problema como um tópico autônomo a ser estudado e mencionou o termo indução apenas duas vezes no livro *Treatise on Human Nature* (1.2.1 e 1.3.7). Apesar disso, considera-se que a ideia moderna tem seu nascedouro num curto argumento desse mesmo livro (1.3.6) (Vickers, 2011), tendo sido posteriormente extendido em *An Enquiry concerning Human Understanding* conforme os exemplos:

> O pão que comi anteriormente nutriu-me. Um corpo de tão sensíveis qualidades foi, àquele tempo, dotado de tais poderes secretos. Mas segue-se que outro pão deverá também nutrir-me em um outro tempo e que qualidades sensíveis similares deverão sempre estar acompanhadas dos mesmos poderes secretos? A consequência não parece de modo algum necessária.[6] (Hume, 2007, 4.2.29, tradução nossa).

Nesse trecho pode-se detectar questionamentos referentes à oposição *Particularidade* × *generalidade*, representada pelo confronto entre um pão específico e os demais pães, bem como a observação da possibilidade de influência do tempo sobre as referidas propriedades (àquele tempo × um outro tempo) – o que chamaremos de *caráter temporal* do problema. Por fim, essas questões são tomadas como objeção à verdade necessária da afirmação de que um outro pão também

[6] The bread, which I formerly eat, nourished me; that is, a body of such sensible qualities was, at that time, endued with such secret powers: but does it follow, that other bread must also nourish me at another time, and that like sensible qualities must always be attended with like secret powers? The consequence seems nowise necessary.

será nutritivo, reforçando o caráter contingente do valor lógico dessa declaração diante da informação disponível. Essa "objeção à verdade necessária" será representada como a oposição *Contingência* $\times$ *necessidade.*

> Quando um homem diz "Eu tenho visto, em todas as ocorrências passadas, tais qualidades sensíveis associadas a tais poderes secretos", e quando ele diz "Qualidades sensíveis similares sempre estarão associadas a poderes secretos similares", ele não é culpado de tautologia, nem são de modo algum essas proposições as mesmas. Dizes que uma proposição é inferida da outra, mas precisas confessar que a inferência não é intuitiva nem demonstrativa. De que natureza é, então? Dizer que é experimental é petição de princípio. Todas as inferências baseadas na experiência supõem, como fundamento, que o futuro é similar ao passado e que poderes similares estarão associados a qualidades sensíveis similares.[7] (Hume, 2007, 4.2.32, tradução nossa).

Essa amostra, embora mais longa que a anterior, concentra-se na questão da circularidade considerada por Weintraub como incluída no argumento geral de Sextus. Aqui, uma vez afirmado que a inferência não é demonstrativa, é explícito o *caráter temporal* por meio da inferência do passado para o futuro. Já o *caráter circular* é derivado do uso da assunção de regularidade da natureza, que equivale justamente a assumir a possibilidade da inferência indutiva em questão. Embora Hume tenha classificado a falácia cometida como sendo **petição de princípio**, atualmente o termo também é entendido como **raciocínio circular** conforme várias consultas, dentre as quais: (Audi, 1999) e (Gensler, 2010, pg. 56).

- Rudolf Carnap – Carnap define o raciocínio indutivo de maneira residual, como sendo todo aquele não dedutível ou não demonstrável, também chamado convencionalmente de raciocínio probabilístico (Carnap, 1962, Preface, V), dando

[7] When a man says, I have found, in all past instances, such sensible qualities conjoined with such secret powers: And when he says, Similar sensible qualities will always be conjoined with similar secret powers, he is not guilty of a tautology, nor are these propositions in any respect the same. You say that the one proposition is an inference from the other. But you must confess that the inference is not intuitive; neither is it demonstrative: of what nature is it, then? To say it is experimental, is begging the question. For all inferences from experience suppose, as their foundation, that the future will resemble the past, and that similar powers will be conjoined with similar sensible qualities.

caráter contingente a todas as sentenças com probabilidade diferente de 0 ou 1. Carnap fala, ainda, que a validade do raciocínio indutivo não depende de pressuposições como a uniformidade do mundo (Carnap, 1962, Preface, V) por estar apoiada em princípios analíticos, que é ponto chave para a componente circular do problema da indução.

Dessa forma, não haveria problema de impossibilidade de justificativa, embora restasse como trabalho a construção de um sistema formal de lógica indutiva. Algumas passagens que contêm os elementos do problema da indução:

> Vamos assumir que um homem X geralmente decide suas ações de acordo com as probabilidades das predições relevantes a respeito das evidências observacionais disponíveis para ele. Esse é um hábito arbitrário ou podemos dar uma justificação para esse modo geral de proceder?[8] (Carnap, 1962, pg. 177, tradução nossa).

Nesse trecho a indução é caracterizada como inferência a partir das "observações disponíveis" (amostra conhecida) para a tomada de decisões em novas situações (o desconhecido). Como tal, implica na assunção da *representatividade amostral* das observações em relação a um domínio de decisões a serem tomadas, cujo questionamento sobre justificação serve de introdução às demais considerações de Carnap sobre as pressuposições da indução (Carnap, 1962, pg. 178–182).

A primeira dessas considerações é justamente a respeito da assunção, "por muitos filósofos" (segundo Carnap), da uniformidade do mundo como pressuposição necessária para a validade das inferências indutivas (inferências probabilísticas) e assim para a justificação da aplicação do método indutivo na determinação de decisões práticas (Carnap, 1962, pg. 178).

Logo em seguida, Carnap considera o *caráter temporal* da uniformidade ao dar um exemplo de inferência indutiva onde está em questão a previsão de chuva para o dia de amanhã, baseada nas informações disponíveis. Por fim, observa o

[8] Let us assume that a man X generally decides his actions in accordance with the probabilities of relevant predictions with respect to the observational evidence available to him. Is this an arbitrary habit or can we give a justification for this general way of procedure?

caráter contingente do resultado previsto e advoga pela relativização do conceito de "uniformidade" de modo a manter a aplicabilidade ao contexto inicial, mas supostamente permitindo o tratamento pelo sistema de lógica indutiva proposto (Carnap, 1962, pg. 178).

- Karl Popper – Em seu livro *The Logic of Scientific Discovery*, considera o sentido usual de inferência indutiva, caracterizando-a como aquela que conduz de enunciados singulares a enunciados universais (Popper, 2002, pg. 3). Em seguida, anuncia a impossibilidade de justificação de uma inferência indutiva baseada na observação de particulares, observando o *caráter contingente* da respectiva conclusão (sempre pode haver um indivíduo que a falseie).

  Dando continuidade à discussão, Popper identifica o problema da indução como a questão de saber se as inferências indutivas se justificam e em que condições[9], também identificando-o como a indagação acerca da validade ou verdade de enunciados universais que encontrem base na experiência, tais como as hipóteses e os sistemas teóricos das ciências empíricas[10].

  Segundo Vickers (Vickers, 2011, section 4.2), Popper considera o próprio método indutivo de aquisição de conhecimento como não importante para a ciência. Consequentemente, todos os problemas inerentes a ele (incluindo o problema da indução) não teriam implicações científicas relevantes, pelo menos do ponto de vista da justificação da escolha entre hipóteses. Mais algumas observações sobre abordagens conectadas à filosofia da ciência serão feitas ao final da listagem.

- Nelson Goodman – em sua abordagem que ficou conhecida como o novo problema da indução (Goodman, 1983, pg. 72), a problemática é analisada a partir de um exemplo informal que confronta duas hipóteses concorrentes e inicialmente equiprováveis, bem como compatíveis com os dados obtidos até então,

[9] The question whether inductive inferences are justified, or under what conditions, is known as the problem of induction.(Popper, 2002, pg. 4)

[10] The problem of induction may also be formulated as the question of the validity or the truth of universal statements which are based on experience, such as the hypotheses and theoretical systems of the empirical sciences.(Popper, 2002, pg. 4)

mas que divergem a partir de determinado tempo t.

Nesse caso, além do questionamento sobre o significado da obtenção de um novo dado para o status de plausibilidade de cada uma das hipóteses, o outro ponto levantado, resgatando o raciocínio de Hume, é a existência de uma divergência entre o senso comum e o raciocínio probabilístico quando aplicado à situação em questão. O problema, também conhecido como Paradoxo de Goodman, pode ser visto na seguinte versão:

> Suponha que no tempo t tenhamos observado várias esmeraldas verdes. Então, temos as declarações evidenciais
>
> *Esmeralda **a** é verde,*
>
> *Esmeralda **b** é verde,*
>
> *etc.*
>
> e essas declarações suportam a generalização:
>
> *Todas as esmeraldas são verdes.*
>
> Mas agora defina o predicado "grue" para aplicar-se a todas as coisas observadas antes de t no caso de serem verdes, e às outras coisas se azuis. Então, também temos as declarações evidenciais
>
> *Esmeralda **a** é grue,*
>
> *Esmeralda **b** é grue,*
>
> *etc.*
>
> e essas evidências suportam a hipótese
>
> *Todas as esmeraldas são grue.*
>
> Logo, as mesmas observações suportam hipóteses incompatíveis a respeito de esmeraldas a serem observadas no futuro; que elas serão verdes e que serão azuis.[11] (Vickers, 2011, 3.2, tradução nossa).

Seguindo com o procedimento de identificação das componentes do problema da indução, pode-se verificar o *caráter temporal* representado pelo uso do tempo t

[11] Suppose that at time t we have observed many emeralds to be green. We thus have evidence statements 'Emerald **a** is green, emerald **b** is green, etc.' and these statements support the generalization: 'All emeralds are green.'. But now define the predicate "grue" to apply to all things observed before t just in case they are green, and to other things just in case they are blue. Then we have also the evidence statements 'Emerald **a** is grue, emerald **b** is grue, etc.' and these evidence statements support the hypothesis 'All emeralds are grue.'. Hence the same observations support incompatible hypotheses about emeralds to be observed in the future; that they will be green and that they will be blue.

como marcador do período de sucesso comum a ambas as hipóteses, a oposição *particularidade* × *generalidade* representada pelo confronto entre as esmeraldas investigadas e o conjunto de todas as esmeraldas e, por fim, o caráter contingente dos valores de verdade das hipóteses concorrentes até o tempo t, diante das informações de que se dispõe.

- Edwin Thompson Jaynes – em seu livro sobre teoria de probabilidades com abordagem bayesiana (Jaynes and Bretthorst, 2003), além de tratar formalmente do uso da indução enquanto lógica de natureza probabilística, fez os seguintes comentários a respeito do problema da indução:

> O problema de "justificar a indução" tem sido difícil para as formulações convencionais da teoria de probabilidades, e também o arqui-inimigo de alguns filósofos, a começar por David Hume (1739, 1777) no século 18. Por exemplo, o filósofo Karl Popper (1974) chegou ao ponto de negar completamente a possibilidade da indução. Ele fez a pergunta retórica: "Estamos racionalmente justificados ao raciocinar a partir de repetidas instâncias sobre que temos experiência para aquelas sobre que ainda não temos experiência?". Isso equivale, literalmente, ao pouco informado robô falando para nós e querendo que respondamos "Não!", mas desejamos mostrar que um robô melhor informado responderá: "Sim, se tivermos informação inicial fornecendo uma conexão lógica entre as diferentes execuções do experimento e dando circunstâncias específicas que possibilitam a indução".[12] (Jaynes and Bretthorst, 2003, pg. 276, tradução nossa).

Se consideramos a citação acima no âmbito do projeto geral de Jaynes, que é usar a teoria de probabilidades como complemento da lógica clássica para

[12] The problem of "justifying induction" has been a difficult one for the conventional formulations of probability theory, and the nemesis of some philosophers beginning with David Hume (1739, 1777) in the 18th century. For example, the philosopher Karl Popper (1974) has gone so far as to flatly deny the possibility of induction. He asked the rhetorical question: "Are we rationally justified in reasoning from repeated instances of which we have experience to instances of which we have no experience?" This is, quite literally, the poorly informed robot speaking to us, and wanting us to answer "No!", but we want to show that a better informed robot will answer: "Yes, if we have prior information providing a logical connection between the different trials" and give specific circumstances that enable induction to be made.

tratar problemas de decisões diante de incerteza, o caráter contingente dos valores de verdade das hipóteses entra quase como definição do processo indutivo considerado, no sentido de que é não-necessário, mas possível, que elas sejam verdadeiras (ou falsas).

Além disso, ao dar a sua resposta à pergunta de Popper, ele acaba por validá-la enquanto representando também a sua visão sobre o problema, com ênfase na questão das instâncias conhecidas enquanto representantes do todo (incluindo as desconhecidas) – a *representatividade amostral*.

- David Charles Stove – Stove tratou de maneira direta o problema em seu livro *The Rationality of Induction*. A partir de textos de Hume e Popper, identificou a "tese cética" sobre a indução como sendo passível de expressão na seguinte forma:

  Para toda evidência e e hipótese h, tais que a inferência de h a partir de e é indutiva, e para qualquer outra informação t tautológica, $\mathrm{P}(h/t.e) = \mathrm{P}(h/t)$.(Stove, 1986, pg. 40)

  O termo $\mathrm{P}(h/t.e)$ lê-se: plausibilidade (ou probabilidade) da hipótese h dadas a evidência e e a informação t.

  É importante observar que o uso do termo "plausibilidade" (qualidade de plausível, aceitável, **razoável**) indica que Stove assume implicitamente a decisão baseada na probabilidade das hipóteses como sendo atitude racional.

  Explicitando esse detalhe, tem-se como consequência da representação acima a irrelevância da evidência e para a razoabilidade (plausibilidade) da adoção da hipótese h, desfazendo-se a base em que se assenta o cálculo de probabilidades condicionadas tão caro à ciência empírica atual.

  Com relação às componentes do problema, embora a formulação probabilística deixe claro tratar-se de inferência de natureza contingente (exceto em casos extremos), a hipótese h pode ser pensada como não tendo nenhuma relação direta com a evidência e além da condicionalização da sua probabilidade. Apesar disso, a abordagem de Stove realizada no capítulo *VI* deixa clara a sua visão da *representatividade amostral* como componente fundamental do problema e, ao mesmo tempo, como base sobre que se desenvolve a solução no decorrer do capítulo.

- Colin Howson – em seu livro especificamente sobre o assunto, chamado *Hume's Problem: Induction and the Justification of Belief*, concorda com a formulação atribuída à Hume, fazendo considerações adicionais das quais foram selecionados os exemplos abaixo:

> O argumento pode ser colocado de maneira possivelmente mais familiar aos ouvidos modernos. Seja P a conjunção de todas as sentenças fatuais conhecidamente verdadeiras. Suponhamos que a inferência a partir de P para uma sentença Q, descrevendo algum evento ainda não conhecido, não é dedutiva (estabelecer que P refere-se a passado e Q a futuro é a primeira parte do argumento de Hume). Segue-se imediatamente da definição de validade dedutiva que em algum subconjunto W de todos mundos possíveis P é verdadeiro e Q é falso. A segunda parte do argumento de Hume pode ser imaginada como originando-se da tentativa de responder à seguinte questão: que informação adicional poderia aumentar a probabilidade de o nosso mundo não estar em W? Bem, o único mundo que conhecemos é o nosso, então presumivelmente a informação deve ser sobre alguns aspecto desse mundo. Mas toda a informação sobre esse mundo que é conhecidamente verdadeira já está em P. Em outras palavras, não há informação adicional.[13] (Howson, 2003, pg. 12, tradução nossa).

Na citação acima, Howson identifica primeiramente o *caráter contingente* da

[13]The argument can be put in a manner possibly more familiar to modern ears. Let P be the conjunction of all factual statements known to be true. Suppose that the inference from P to a statement Q describing some event not known to be true is not deductive (establishing that this is so where P stands for 'past' and Q for 'future' is the first part of Hume's argument). It follows immediately from the definition of deductive validity that in some subset W of all the possible worlds in which, like ours, P is true and Q is false. The second part of Hume's argument can be imagined as arising from trying to answer the question: what further information could be appealed to which would make it more likely that our world is not in W? Well, the only world we know is this one, so the information must presumably be about some aspect of this world. But the only information we have about this world that is known to be true is already in P. In other words, there is no further information.

inferência ao considerá-la como não dedutiva. Em seguida, o *caráter temporal* surge a partir da referência ao argumento de Hume sobre inferências do passado para o futuro, embora subentenda-se do texto que as inferências temporais são um caso especial das não-dedutivas.

Por fim, ao confrontar o "nosso" mundo particular P com todos os mundos possíveis ou especificamente com o subconjunto W, é posta a oposição entre *particularidade e generalidade* com relação às conclusões com base no conhecimento disponível.

Além da identificação de componentes do problema, Howson ressalta a gravidade das objeções de Hume argumentando que elas aplicam-se igualmente às tentativas de justificação das inferências indutivas enquanto "prováveis"[14], fenômeno que possibilita a criação da ilusão de que o problema da indução tem como principal obstáculo à ciência a simples impossibilidade da certeza quanto à verdade da inferência.

Tal fosse o caso, uma vez aceita a incerteza como inevitável em praticamente todos os problemas científicos, a decisão baseada na probabilidade seria suficiente para reafirmar a racionalidade da indução, o que não pode ser conseguido assumindo-se, como Howson mostra, que o problema atinge todas as inferências não-dedutivas.

Algumas outras abordagens também foram feitas tendo como pano de fundo a evolução da ciência, como o fez Popper, mas com opinião diferente com relação à resolução do problema. Nessas abordagens, segundo Hempel (Hempel, 1981), o problema da decisão entre hipóteses concorrentes diante das evidências experimentais (uma das principais consequências do problema da indução) é deslocado para o âmbito geral da compatibilidade e relacionamento com as outras hipóteses que compõem o conjunto do conhecimento científico do momento em questão.

[14]It is the "going farther" that is the original, and if correct quite devastating, part of Hume's argument. That there is no deductive link between statements about past and future had been known since antiquity. Where Hume goes beyond the traditional sceptical position is in arguing the link cannot in principle justified even as a "probable inference".(Howson, 2003, pg. 12, tradução nossa)

Desse novo ponto de vista, critérios de decisão alternativos – pode-se pensá-los como em escala macroscópica – foram propostos notadamente por Kuhn (simplicidade, precisão, escopo, etc.), Laudan (número de problemas importantes resolvíveis) e Richard Rudner (relevância dos possíveis erros decorrentes da aceitação da hipótese) (Hempel, 1981). Como resultado, embora permaneça a questão da relação entre hipótese e evidência enquanto parâmetro de decisão, ela perde importância quando comparada aos outros critérios e, por isso, tais abordagens serão equiparadas à de Popper quanto à importância do problema, embora divirjam quanto ao papel da indução na ciência.

Teorema *No Free Lunch*

Além das amostras anteriores, abordagens probabilísticas recentes (1996 e 1997) nos campos de "aprendizagem de máquina" e algoritmos de busca e otimização consistem em uma formalização computacional de método indutivo capaz de fornecer um modelo a que se aplicam objeções análogas às de Hume, no sentido de que não é possível a obtenção de conhecimento livre de vieses, bem como a obtenção de certezas a respeito do futuro a partir de amostras passadas. O resultado é resumido pelo jargão *There's No Free Lunch*, derivado de um ditado popular[15] que transmite a ideia fundamental segundo a qual é impossível obter algo sem custos.

Pode-se entender essa ideia original como aplicada à formalização no sentido de que, considerando duas regras de fazer predição e um conjunto de universos possíveis, a vantagem de uma das regras sobre a outra em determinado universo será compensada por uma desvantagem em um outro. Ou seja: a boa performance inicial (*free lunch*) será "paga" em forma de uma má performance futura. Um exemplo simples é fornecido em (Forster, 2005).

Optou-se por colocar essa amostra legítima do problema como um item distinto devido a ela ser recente e proveniente de autores que não estariam cobertos pelo critério de importância filosófica utilizado, mas, ao mesmo tempo, representar uma

[15] Usos da frase já foram encontrados em textos de 1930 e 1940, mas a origem exata ainda é incerta. (Safire, 1993)

abordagem mais rigorosa do ponto de vista do formalismo matemático. Indo direto às conclusões passíveis de serem extraídas pelo modelo, tem-se:

"NFL para aprendizado supervisionado formaliza Hume: a ciência não pode dar garantias a respeito de experimentos futuros baseada em resultados de experimentos passados."[16] (Wolpert, 2012, tradução nossa).

e

"(Wolpert, 1996) mostra que em um cenário sem ruído onde a função-perda é a taxa de erro de classificação, caso se está interessado no erro extra situação de teste, não há distinção a priori entre algoritmos de aprendizado."[17] (Sewell, 2012, tradução nossa).

No caso acima, o termo *a priori* está sendo usado apenas temporalmente em relação aos dados a serem adquiridos. Ou seja: qualquer informação disponível antes da aquisição de dados, é considerada informação *a priori*. Essa será a definição considerada nesse estudo, reforçando que o tempo usado para a caracterização é aquele relativo à inferência em questão. Assim, é possível que o que foi considerado dado para uma pessoa, passe a ser informação *a priori* para outra (Jaynes and Bretthorst, 2003, pg. 87).

Ao afirmar, na primeira citação, que *a ciência não pode dar garantias* (oposição *contingência* × *necessidade*), o autor parece não considerar uma possível justificação probabilística como incluída no problema da indução, na qual se abriria mão da certeza da inferência mantendo a racionalidade por meio da escolha da opção mais provável, atualizada à medida que mais dados tornam-se disponíveis, ainda que inicialmente as hipóteses tenham começado com a mesma atribuição de probabilidade (por exemplo, via princípio da indiferença). A despeito disso, são claras as componentes relativas à regularidade do universo, bem como a impossibilidade de distinção entre induções boas ou más.

[16] NFL for supervised learning formalizes Hume: Science cannot give guarantees about future experiments based on results of previous experiments.

[17] (Wolpert, 1996) shows that in a noise-free scenario where the loss function is the misclassification rate, if one is interested in off-training-set error, then there are no a priori distinctions between learning algorithms.

Conhecidos os representantes legítimos, faz-se conveniente a identificação de falsas instâncias postas como exceções por não conterem alguma componente vital da ideia genérica sobre o problema.

Exceções

- Indução matemática – apesar do nome, a indução matemática é, na verdade, um método dedutivamente válido de inferência que possibilita conclusões genéricas a partir de uma prova finita, sempre que o conjunto-objeto em questão puder ser expresso por uma definição indutiva.

- Inferência a partir de "lei" e amostra – considerando-se um conjunto definido por características condicionais, mas de aplicação geral, a observação da característica de um dos membros seria suficiente para a inferência da mesma característica para os demais.

  Ex.:

  Seja um mundo formado por blocos que podem ter formas variadas, mas que obedecem à seguinte lei: "todos os blocos têm a mesma forma". Ao ser observado um bloco cúbico, a inferência de que todos os demais são cubos não tem o caráter contingente intrínseco à ideia de indução, já que a conclusão segue necessariamente da premissa inicial e da observação.

Apesar de diferentes, as exceções mostradas têm em comum a não adequação ao critério da contingência, necessário para a caracterização de uma inferência como indutiva, conforme as considerações anteriores.

Voltando aos exemplos considerados como representantes do problema da indução, pode-se construir uma tabela contendo as principais componentes identificadas nos comentários relativos às citações e demais considerações (em sua maior parte destacadas em itálico), resultando em:

Com base nas observações anteriores e particularmente na tabela acima, estamos em condições de concluir o processo de explicação de forma a facilitar a delimitação

Tabela 1.1: Características das amostras representativas do problema

	Caráter temporal	Caráter circular	Contingência × necessidade	Particularidade × generalidade	Representatividade amostral
Sextus		√	√	√	
Hume	√	√	√	√	
Carnap	√		√		√
Popper			√	√	
Jaynes			√		√
Stove			√		√
Howson	√		√	√	
Wolpert	√		√	√	√
Goodman	√		√	√	

do que será considerado como "o" problema a ser tratado, conforme se segue.

1.3.3 *Explicatum*

A primeira observação a ser feita diz respeito às duas últimas colunas da tabela: *Particularidade × generalidade* e *Representatividade amostral*.

O termo *representatividade amostral* está sendo tomado no sentido de equivalência aproximada[18] entre amostra e população total com relação à característica considerada, de forma que uma inferência com base na amostra também seria válida (com uma margem de erro esperada) para a população.

A partir desse entendimento, essas componentes podem ser consideradas como equivalentes no que diz respeito ao seus significados para o problema da indução. Tal equivalência pode ser notada facilmente se considerarmos que uma amostra pode ser vista como um particular numa população de amostras possíveis (o *universo* de amostras possíveis) e que as questões sobre a equivalência entre amostra e população também são perfeitamente aplicáveis em relação a particular e universal.

Se passarmos agora a considerar o papel das componentes restantes nas diversas abordagens do problema representadas pelas amostras, veremos que tanto o *caráter temporal* quanto a questão representada pela oposição *Particularidade × generalidade*

[18] Os critérios que definem a maior ou menor aproximação, bem como o próprio tamanho e maneira de se construir a amostra, não são objeto desse estudo.

(essa última tomada agora como abarcando também o problema da *representatividade amostral*), são usadas basicamente para afirmar a incerteza quanto ao valor de verdade da inferência indutiva.

Tal manobra pode ser observada, por exemplo, na afirmação de que o futuro não é necessariamente igual ao passado (relativamente ao *caráter temporal*), assim como ao ser afirmada a possibilidade de que um particular ainda desconhecido se contraponha à generalização[19]. Vistas dessa forma, portanto, ambas as componentes seriam casos especiais da problemática representada pela oposição *contingência* $\times$ *necessidade*.

Uma vez tomada a contingência como representando o caráter problemático na base das componentes consideradas até então, resta investigar a importância da questão da circularidade para o tratamento da indução e de suas implicações. Eis que enquanto a contingência[20] é característica do processo de inferência indutivo e, ao mesmo tempo, obstáculo à sua justificação, a circularidade é característica não da indução, mas de algumas estratégias de justificação.

Diante dessas observações, um modelo finito representando um processo de inferência indutivo (entendido como aquele em que a inferência não é certa) permite objeções análogas às observadas nas abordagens escolhidas como representantes do prolema da indução.

Por fim, o modelo de inferência indutiva pode ser descrito conforme abaixo:

"Considere um conjunto $\boldsymbol{A}$ *sobre que conhecemos apenas alguns elementos. Ou seja: há um subconjunto próprio de* $\boldsymbol{A}$*, que podemos chamar de* $\boldsymbol{CA}$*, de elementos que sabemos pertencer a* $\boldsymbol{A}$*. Além disso, há outros elementos pertencentes ao domínio de* $\boldsymbol{A}$*, mas que estão fora de* $\boldsymbol{CA}$ *e cuja pertinência a* $\boldsymbol{A}$ *nos é desconhecida. Nesse contexto, a inferência indutiva é a predição de que um dado elemento* $\boldsymbol{b}$*, fora de* $\boldsymbol{CA}$*, pertence ou não a* $\boldsymbol{A}$*, usando como informação apenas o que conhecemos sobre* $\boldsymbol{CA}$*."*

Percebe-se que enquanto a modelagem mais geral do problema da indução seria

[19] O que significaria dizer que a amostra não seria representativa em função de um ou mais elementos, ainda desconhecidos, da população.

[20] Mais uma vez: contingência em relação ao valor de verdade da inferência, cuja informação disponível permite conceber tanto como verdadeiro quanto falso. Também entendida como representando a incerteza da inferência.

representada pela tentativa de encontrar a "função característica"[21] de A a partir de uma amostra de A (no caso, CA), o modelo acima enfatiza a predição individual.

Isso pode ser justificado se considerarmos a possibilidade de o agente em questão poder rever sua inferência sobre a população geral a cada amostra obtida. Além disso, o que se espera para a próxima amostra tem relação direta com a natureza dessa população geral e também possui o caráter de incerteza considerado componente fundamental do problema. Tal ênfase na próxima amostra como representante do processo de indução – chamado de *singular predictive inference*[22] – também é partilhada por Mill e Carnap (Vickers, 2011).

Assumindo esse modelo restrito como domínio de estudo, as seguintes questões a respeito do processo de inferência são pertinentes:

1. O que é predição e qual sua natureza?
2. É possível predição sem amostra representativa?
3. Dada uma amostra, como classificá-la como representativa?
4. Como corrigir o status da amostra, com relação à representatividade, dado o acréscimo de novos elementos?
5. Como escolher entre diferentes possibilidades de predições, dado um conjunto de amostras e algoritmos de predição?

Com exceção dos dois primeiros itens, pode-se identificar instâncias dessas perguntas nos exemplos representativos listados anteriormente. Por exemplo:

- a consideração de uma amostra como representativa pode ser entendida como a assunção de que o passado (representado pela amostra) é similar ao futuro (próximos elementos a serem "sorteados");
- as previsões possíveis são equivalentes às saída de diferentes algoritmos preditores, dadas as amostras anteriores;

[21] Entendida como a função que retorna, para cada elemento do domínio, a pertinência em relação ao conjunto A. (Jech and Hrbacek, 1999)

[22] Inferência preditiva singular

- a própria questão de escolher entre algoritmos preditores ecoa a pergunta de Hume sobre como distinguir uma indução boa de uma má.

Dessa forma, será considerado nesse estudo que as respostas às questões referentes ao processo de aquisição de conhecimento simplificado podem ser compreendidas, até onde a equivalência permite, como aplicáveis ao problema da indução conforme percebido por Hume e demais estudiosos que trataram do tema.

Capítulo 2

As implicações e o modelo

"O Raciocínio Indutivo, que há muito tem sido a glória da ciência, terá deixado de ser o escândalo da filosofia?"[1] (Broad, 1952, pg. 142, tradução nossa).

A frase acima expressa, de maneira extremamente concisa, a importância das implicações do problema da indução sobre o processo de aquisição de conhecimento científico sobre o mundo, sendo essa a própria fonte da sua característica problemática, bem como da consequente necessidade de resolução.

Uma vez especificado de maneira mais precisa e limitada, mais de acordo com as pretensões dessa análise, faz-se prudente a checagem da existência da equivalência das implicações "originais" àquelas observáveis entre as questões escolhidas para representar o problema e o que seria o processo de aquisição de conhecimento em nosso modelo.

Dessa forma, são propostas as seguintes equivalências:

2.1 O desmoronamento do conhecimento científico

É impossível, hoje em dia, não observar a importância que métodos indutivos têm na obtenção do conhecimento científico, bem como a nítida distinção entre o status de autoridade daquilo que pode usar o adjetivo científico e do que é tratado apenas como

[1]Inductive Reasoning, which has long been the glory of Science, will have ceased to be the scandal of Philosophy?

conhecimento “informal” ou, pior, pseudo-ciência.

A autoridade obtida pela ciência, reforçada pelo grande volume de resultados que ela demonstra, passou a ser cobiçada por qualquer crença que pretenda ser aceita e divulgada a um grande número de pessoas, a não ser, talvez, aquelas cujo único fim é trazer a paz interior sem implicações relacionadas ao mundo externo, como é o caso de algumas religiões.

Considerando os meios a partir dos quais uma hipótese pode galgar mais um degrau em direção ao tão cobiçado status de científica, há uma assimetria interessante que se observa em relação a esses “caminhos”: o dedutivo, embora seja o mais rigoroso, parece ser muito mais difícil de trilhar e, por isso, dá origem a menos hipóteses aceitas; enquanto isso, o caminho indutivo virou uma verdadeira via-expressa e, mesmo não sendo tão rigoroso, goza praticamente das mesmas vantagens em termos de status, o que foi conquistado com o estabelecimento de inúmeros procedimentos metódicos e estatísticos visando a diminuição de ambos os tipos de erros, conhecidos como Tipo 1 e Tipo 2 (falso negativo e falso positivo, respectivamente).

Devido a essa discrepância na quantidade de “verdades” originadas de cada um desses caminhos, não é exagero dizer, quando se fala em ciência enquanto conhecimento justificado e aceito pela maioria, que se está a fazer referência principalmente ao conhecimento de “origem indutiva”.

É justamente da aparente indissociabilidade entre ciência e justificação - e da consequente constatação da ausência dessa última - que o escândalo mencionado por C. D. Broad torna-se relevante, fazendo com que cientistas e principalmente filósofos se vejam forçados a conviver com a agonia da sensação de iminente desmoronamento do conhecimento indutivo, ou sintam-se no dever de providenciar tal justificativa.

Passando à análise do modelo proposto, é necessário o estabelecimento do que corresponderia ao conhecimento e o que seria o processo indutivo responsável por obtê-lo. Assim, podemos ver que impacto teriam as respostas às nossas questões, já supostas representantes do problema da indução, no conhecimento adquirido e no próprio processo de aquisição, comparando-o, em seguida, ao escândalo mencionado anteriormente.

Conhecimento

Se entendermos o conhecimento no sentido de crença justificada em uma alegação sobre o mundo (conhecimento empírico), ele englobaria tanto as amostras passadas, com "grau de certeza" igual a 100% (o que corresponderia à probabilidade de amostrar um elemento e dado que foi amostrado e), quanto as predições feitas usando o processo de aquisição "científico" (o algoritmo preditivo) e o respectivo "grau de certeza" menor que 100%, já que a predição pode não se confirmar (é contingente).

Processo de aquisição

O processo de aquisição abordado consiste no conjunto de regras usadas para amostrar, estabelecer a predição, confirmá-la (por meio de comparação entre predição e a amostragem efetiva), calcular a probabilidade de acerto, bem como quaisquer outras informações diferentes da caracterização dos próprios elementos, desde que usados direta ou indiretamente para adicionar novas amostras ao conjunto amostragem ou para predizer amostras futuras.

Diante das equivalências estabelecidas e das instâncias dos problemas consideradas no capítulo anterior, as seguintes implicações são observadas:

- **Ausência de conhecimento** - se a amostra não é representativa, não pode ser compreendida como conhecimento sobre o objeto-alvo (o conjunto universo desconhecido), no sentido de poder ser usado para extrapolações a seu respeito, sejam elas características do universo como um todo ou simplesmente relativas à predição do próximo elemento a ser amostrado, ainda que haja algoritmos preditores disponíveis.

- **Algoritmos preditores** - supondo a existência de algoritmos preditores, independentemente de serem eles considerados como parte do próprio conhecimento, da ausência de justificação da indução e da consequente impossibilidade de escolha entre esses algoritmos, ou mesmo de distinção entre eles e uma escolha aleatória à longo prazo, ecoa o lema: *there's no free lunch*, no sentido de que todos os algoritmos preditivos são equivalentes à longo prazo.

É importante ressaltar que a capacidade de predição implica a assunção inicial de um nível mínimo de representatividade amostral a partir do qual seriam inferidos os próximos elementos a serem obtidos, embora não implique na necessidade da correção dessa assunção nem na possibilidade de certeza sobre a inferência em questão.

2.2 A impossibilidade do "conhecimento probabilístico"

Segundo David Stove, enquanto para Salmon há a questão de por que se deve "apostar" na hipótese mais provável, para Hume não há sequer hipótese mais provável (não há conexão entre passado e futuro e, assim, a probabilidade de uma dada hipótese é a mesma com ou sem dados). (Stove, 1986, pg. 193)

Ao reformular essa ideia em termos probabilísticos, ressultando na equação $\mathrm{P}(h/t.e) = \mathrm{P}(h/t)$ explicada anteriormente em 1.3.2, Stove torna claro o significado de tal assunção sobre o processo indutivo de aquisição de conhecimento, além de facilitar pensarmos sobre a que tipo de modelo essa formulação seria aplicável.

Do ponto de vista de um agente com o propósito de maximizar os acertos de predições, tal formulação probabilística implica na não existência de qualquer tipo de argumento racional a favor de qualquer algoritmo preditor, o que consistiria em algo além do mero "hábito" (nas palavras de Hume), considerado como única fonte de explicação de tal escolha.

Assim, deve-se observar a ausência de qualquer ligação cognitiva entre os cálculos probabilísticos, independentemente do nome dado a eles, e a escolha dos algoritmos preditores.

2.3 A Modelagem e as Ferramentas

Até agora têm sido usados termos probabilísticos e de teoria de conjuntos sem muita justificativa ou explicação adicional, mas de forma que pudessem ficar evidenciados seu papel e importância nessa abordagem.

Uma vez apresentados os principais conceitos e as explicações cabíveis, pode-se escolher uma situação-exemplo prática a ser simplificada e a partir da qual será construído o modelo-referência do processo de aquisição de conhecimento, de forma a dar suporte às ideias expostas até então, bem como ao seu desenvolvimento e detalhamento necessários.

2.3.1 O exemplo prático: lançamento de moeda

Suponhamos que um cidadão chamado Wittinho está lendo tranquilamente em uma praça. Após algum tempo, senta-se ao lado dele um cara que se apresenta como Sr. Moliko, de posse de uma moeda, que lança constantemente para o alto checando o resultado e dizendo-o em voz alta: cara, coroa, coroa, coroa, cara, coroa, e assim, incessantemente.

Vendo esse comportamento curioso, Wittinho pergunta se o Sr. Moliko está fazendo algum tipo de experimento, quando Moliko responde que apenas acha o comportamento daquela moeda específica muito estranho, pois ela tende a favorecê-lo em apostas, mesmo quando outra pessoa a lança e tem a liberdade de fazer a escolha primeiro.

Wittinho, sem acreditar na história, pede para eles apostarem um pouco, desde que o valor da aposta seja estabelecido por ele.

Diante dessas considerações, as possibilidades *relevantes* que passam pela cabeça de Wittinho são 3:

1. A moeda é justa;
2. A moeda tem viés para cara;
3. A moeda tem viés para coroa.

Como Wittinho poderia fazer os lançamentos, foi descartada a possibilidade de a moeda ser trocada por Moliko a cada jogada, conforme a aposta, de modo a beneficiá-lo. Sendo assim, a natureza da moeda seria mantida a mesma e de acordo com exatamente uma das alternativas pensadas.

Se quiséssemos resguardar ainda mais Wittinho com relação à capacidade de lidar com o experimento, poderíamos imaginar que ele teria em mente uma última hipótese que atribuiria algum controle misterioso de Moliko sobre o resultado da moeda, ainda que a crença nessa possibilidade fosse praticamente desprezível. O resultado de tal procedimento é que, a não ser frente a dados bastante inesperados (por exemplo: no caso de Moliko realmente ganhar todas as apostas) esta 4ª hipótese permaneceria muito menos plausível que qualquer uma das *relevantes*.

Nos casos onde efetivamente ocorrem tais resultados extremos, teríamos o fenômeno da "ressurreição de hipóteses" (Jaynes and Bretthorst, 2003, pg. 105), que apesar de importante para os testes de hipóteses em geral, pode ser desconsiderado sem prejuízos, dados os objetivos aqui pretendidos. Dessa forma, as hipóteses ditas relevantes serão consideradas como sendo as únicas possíveis e mutuamente excludentes.

No caso 1, em que a moeda é justa, Wittinho espera que não haja ganho significativo para qualquer um dos apostadores, excluindo qualquer possibilidade de controle que ele mesmo possa exercer sobre o resultado do lançamento. Como não há aposta privilegiada que se possa fazer, já que tanto cara quanto coroa teriam probabilidades de 50% de ocorrência e seriam eventos independentes (não havendo relação entre ocorrência de eventos passados e eventos futuros), ele estaria disposto a correr o risco de perder um pouco de dinheiro.

Em qualquer dos outros casos, entretanto, ele imagina que ao longo dos lançamentos conseguiria definir com uma probabilidade crescente de acerto o viés da moeda e, a partir daí, usar essa informação para ganhar de Moliko nas jogadas restantes, compensando as eventuais perdas que sofreria até então.

Como não dispõe de muito tempo, Wittinho resolve estabelecer o limite de 10 lançamentos para o experimento, quer tenha obtido saldo positivo ou negativo após atingido esse limite. Imaginemos, para fins ilustrativos, que haja apenas a possibilidade de viés extremo: 100%. Dessa forma, se a moeda não fosse justa, os lançamentos resultariam em uma sequência de 10 caras ou 10 coroas, conforme o caso. Esse viés extremo também pode ser entendido como sendo equivalente a uma situação onde as moedas não-justas teriam duas faces iguais (duas caras ou duas coroas), mas em que o experimento fosse realizado de forma que o jogador fizesse o lançamento e observasse o resultado sem ter como examinar e identificar a moeda.

É importante observar que a perceptibilidade do viés da moeda dependerá tanto do número de lançamentos a serem executados quanto de sua magnitude, que pode ser representada pela diferença entre as probabilidades reais de ocorrência dos eventos e aquelas esperadas caso a moeda fosse *honesta*. Por exemplo: para uma moeda cuja probabilidade real de obtenção de cara é de 3/4, diz-se que há um viés de magnitude 1/4 a favor dessa ocorrência (50% em relação a 1/2).

Como inicialmente todas as hipóteses são igualmente plausíveis e ainda não foi realizado nenhum lançamento, a primeira aposta deverá ser feita de maneira aleatória no sentido de que não há justificativa baseada no objetivo imediato, que é acertar o resultado do primeiro lançamento, mas apenas para dar início ao processo de aquisição de conhecimento.

Diante dessa restrição poderíamos pensar nos seguintes casos sob que Wittinho poderia raciocinar, supondo-se que ele fez a primeira aposta em coroa:

Caso 1 - primeiro lançamento resulta em cara

Nesse caso, dadas as considerações anteriores, estaria eliminada automaticamente a moeda com viés de 100% para coroa, restanto as hipóteses:

a) A moeda é justa;

b) A moeda tem viés de 100% para cara;

Se nos limitarmos apenas a essas informações, já seria possível escolher uma das hipóteses restantes em detrimento da outra, mesmo que possamos mudar a escolha à medida que mais resultados sejam obtidos? A partir dessa escolha, já poderíamos tentar acertar o próximo resultado considerando o que se esperaria como implicação dela?

Sem entrar em detalhes quantitativos, reservados ao tratamento posterior, algumas considerações qualitativas são capazes de indicar que princípios seriam usados como guias no raciocínio sobre a decisão e, tratando-se particularmente desse caso ilustrativo, até mesmo chegar a efetuá-la.

A primeira consideração a ser feita diz respeito às probabilidades iniciais ou *prior probabilities* que Wittinho atribuiria a cada uma das hipóteses restantes. Apesar da aparente simplicidade da tarefa, o fato de devermos levar em conta toda a informação relevante disponível para o agente, inclusive *prior information* (ou simplesmente *priors*), faz com que esse passo seja merecedor de atenção especial. Por *prior* entenda-se qualquer informação não considerada inicialmente como dado do problema, conforme explicado anteriormente em 1.3.2.

Assim, pode-se imaginar estados cognoscíveis que vão desde a ausência total de *priors*, o que levaria à atribuição de probabilidades iguais aos eventos, segundo o que se tornou conhecido como *princípio da indiferença* (Keynes, 1921), a toda uma hierarquia crescente de sabedoria sobre o mundo no que se relaciona à situação específica abordada.

Seguem alguns exemplos da diversidade da natureza de tais informações:

I - Conhecimento do comportamento de moedas examinadas anteriormente;

II - Crença sobre probabilidade de encontrar uma moeda *viciada*;

III - Crença sobre comportamento esperado de moedas viciadas;

IV - Conhecimento sobre situações similares que se revelaram golpes de malandros;

V - Crença sobre reputação de Moliko com relação à honestidade;

VI - Crença na impossibilidade de Moliko influenciar o resultado nas condições escolhidas;

VII - Crença sobre reputação de Moliko com relação à sanidade, etc.

Una-se a isso o fato de que cada uma dessas informações pode ser concebida como a conclusão de um argumento indutivo anterior, herdando a incerteza que o caracteriza e transferindo-a adiante na cadeia de inferências. É impressionante que o nosso cérebro corriqueiramente lide com esse tipo de problema e consiga atribuir as probabilidades *a priori* às afirmações e hipóteses consideradas, numa demonstração de potencial que não é ofuscada mesmo quando o processo é realizado de maneira não-ótima.

Para manter a simplicidade, suponhamos que a informação *a priori*, nesse caso, tenha sido usada apenas para se chegar às possibilidades consideradas, mas que não haja nenhuma distinção entre elas do ponto de vista de plausibilidades, de modo que Wittinho ainda estaria igualmente indeciso quanto ao caso de a moeda ser justa ou ser viciada. Ainda assim, há pelo menos duas maneiras de raciocinar para escolher em qual hipótese apostar e, baseado nela, que resultado deve-se escolher para o próximo lançamento.

A primeira maneira diz respeito ao grau de confirmação que o resultado obtido representa para cada uma das hipóteses. Apesar de ambas serem compatíveis com a ocorrência considerada, suas plausibilidades são afetadas diferentemente, conforme pretende esclarecer a explicação a seguir.

Imaginemos que as moedas em questão são retiradas de uma caixa contendo apenas justas ou com viés de 100% para cara. A extensão do princípio da indiferença a esse nível de análise levaria a concluir que, dando continuidade aos sorteios, iríamos obter aproximadamente o mesmo número de viciadas e de honestas, já que qualquer expectativa de resultado diferente estaria em contradição com a nossa assunção inicial de probabilidades iguais.

Apesar de esperarmos obtê-las em mesma quantidade, enquanto o lançamento das justas tenderá a resultar em um número igual de caras e coroas (metade dos lançamentos), as viciadas resultarão sempre em caras, fazendo com que o número de caras obtidas por meio de moedas viciadas seja o dobro daquelas obtidas com as honestas.

Assim, como resultado dos lançamentos (apenas uma vez para cada moeda sorteada), se formássemos um conjunto de moedas que deram cara, ele tenderia a ser composto por 2/3 de viciadas e de 1/3 de honestas, à medida que seu número de elementos aumentasse. Desse ponto de vista, fica claro que o conhecimento de que uma moeda pertence ao conjunto (resultou em cara) é favorável à hipótese de que se trata de uma viciada, já que elas são a maioria.

Pensando de uma outra maneira, mais centrada na otimização de resultados, pode-se simplesmente decidir por apostar novamente em cara baseado no fato de que se a moeda for justa esse resultado é tão provável quanto coroa, independentemente de resultados anteriores. Por outro lado, se for verdade que a moeda tem viés, a chance de acerto é de 100%, o que faz com que a aposta em cara seja ótima no âmbito geral.

Dessa forma, enquanto persistir a ocorrência de cara como resultado, ambos os raciocínios anteriores levam à mesma previsão para o próximo lançamento.

Caso 2 - primeiro lançamento resulta em coroa

O raciocínio anterior aplica-se igualmente a esse caso, com a diferença de que agora as hipóteses restantes seriam:

a) A moeda é justa;

b) A moeda tem viés de 100% para coroa;

Seguindo as considerações mostradas no caso 1, mas adaptando-as à ocorrência de coroa, chegaríamos à conclusão de que esse resultado favoreceria a hipótese de viés da moeda (para coroa), em detrimento da que propõe tratar-se de uma moeda justa. Da mesma forma, o *raciocínio pragmático* visando o acerto no próximo lançamento indica como sensata a manutenção da aposta em coroa.

Avançando um pouco no experimento, suponhamos que até o 9º lançamento, para usar um caso extremo, obteve-se apenas coroa como resultado. Embora não se possa eliminar a possibilidade de se estar diante de um caso raro de ocorrência para uma moeda justa, a essa altura Wittinho estaria muito mais confiante na existência do viés que na *honestidade* da moeda, o que seria uma atitude racional desde que ambas as hipóteses tenham iniciado com a mesma plausibilidade atribuída, como no caso 1, pela aplicação do princípio da indiferença.

Devido a essa confiança crescente no viés, suponhamos que a última aposta também tenha sido em coroa, mas que o resultado do último lançamento tenha sido cara, indicando, no fim das contas, que se tratava de uma moeda honesta. Curiosamente esse último resultado levanta vários questionamentos, independentemente de não significar muito do ponto de vista geral (afinal de contas obteve-se 9 acertos), por exemplo: qual a razoabilidade de se obter predições acertadas usando hipóteses falsas? O que está por trás da dificuldade intuitiva em lidar com eventos raros e de que forma a teoria de probabilidades pode ajudar?

Uma dessas falhas na intuição, conhecida como falácia do apostador (Lehrer, 2009), teria coincidentemente obtido sucesso na história escolhida para o experimento (9 coroas e 1 cara), se considerarmos justamente a última aposta.

Explicando melhor: é comum a ocorrência desse fenômeno intuitivo que nos leva a desconsiderar a independência entre eventos e atribuir uma probabilidade menor a eventos com maior número de ocorrências passadas e, inversamente, maior àqueles que ocorreram menos. Como consequência, a ocorrência de tal série de coroas fortaleceria a crença em que, para o próximo lançamento, o resultado cara seria mais provável. Ainda segundo a crença, essa probabilidade aumentaria proporcionalmente ao tamanho da sequência.

Feitas essas considerações, é importante destacar que é perfeitamente possível que uma outra pessoa, mesmo diante das 9 coroas, ainda atribuísse uma probabilidade maior à hipótese de tratar-se de uma moeda justa e, assim, pudesse estar justificada em apostar em cara, esperando uma probabilidade de acerto de 50%.

Essa atitude divergente frente aos mesmos dados seria explicada justamente pela diferença na atribuição das probabilidades às hipóteses anteriormente ao primeiro lançamento, especificamente quanto ao item II da lista de possíveis informações *a priori* (em 2.3.1), o que levaria a uma assimetria com relação à quantidade de evidência necessária para que se considere uma delas preferivelmente à outra.

Em outras palavras, a pessoa que continuasse apostando na honestidade da moeda teria iniciado a sequência de lançamentos com uma atribuição de probabilidades diferentes para as hipóteses, resultado de um raciocínio onde outros fatores, que não o princípio da indiferença, exerceram maior influência.

Diante do que foi visto até aqui, espera-se que tenham sido explicitados os princípios básicos que possibilitariam a Wittinho, mesmo dadas as restrições impostas, escolher entre hipóteses e realizar as predições.

Além disso, é possível notar a importância da informação a priori tanto como ponto de partida para o processo de *aprendizado*, quanto como explicação para as diferenças aparentemente inexplicáveis entre as escolhas de agentes diversos que usem estritamente o mesmo processo de inferência proposto.

2.3.2 Modelagem

Em linha com as informações dadas inicialmente em 1.1, e tendo em vista o exemplo da moeda que acabou de ser mostrado, a escolha do modelo a ser usado na análise do problema da indução foi guiada pela observação de que ele se relaciona particularmente com o processo de aquisição de conhecimento científico – doravante chamado abreviadamente de **aprendizado**.

Do ponto de vista de Wittinho, o aprendiz considerado na situação hipotética, esse processo resultaria na atualização das plausibilidades atribuídas às hipóteses consideradas, de acordo com a aquisição de mais informação representada pelo conhecimento dos resultados dos lançamentos.

Dessa forma, buscou-se construir uma representação o mais simples possível, mas que ainda preservasse a essência dos principais conceitos usados na caracterização do processo, bem como do problema que atua sobre ele: o **problema da indução**.

Diante dessas considerações, um ponto de partida razoável é a identificação dos participantes do processo de aprendizado, que são:

- um agente que busca o conhecimento: chamemo-lo de *sujeito* (Wittinho);

- algo de onde o conhecimento é obtido: chamemo-lo de *experimento*;
- e, claro, o objeto da busca: o *conhecimento* (resultados dos lançamentos).

Independentemente do tipo do objeto da busca, e de que processos de aquisição seriam considerados legítimos (pontos a serem tratados mais à frente), há ainda um conceito sem o qual não é possível representar a problemática a ser estudada: a *predição*.

Realizar uma predição, no sentido aqui considerado, é nada mais que especular sobre um resultado futuro a partir do que já se conhece. No caso do exemplo fornecido, a predição é justamente a aposta feita por Wittinho para o próximo lançamento, dado o estado cognoscível total atual, composto pela informação inicial e pelo conhecimento dos resultados dos lançamentos já efetuados.

Pode-se atribuir um caráter racional ao ato de tentar uma predição sempre que se observar que o agente em questão obtém alguma vantagem caso consiga antecipar o próximo resultado. No âmbito geral, a justificativa depende de determinado balanço entre a taxa de acerto, o "custo" da aposta e o "prêmio" oferecido, o que pode ser visualizado de maneira mais clara pelo cálculo do valor esperado para cada ação disponível para o agente.

Na situação imaginada, onde o prêmio é em dinheiro, as decisões possíveis são: não apostar, apostar em cara ou apostar em coroa.

Para os fins pretendidos, basta considerarmos que o valor esperado "visto" por Wittinho é positivo, o que é coerente com as hipóteses que ele considerou como relevantes (com as respectivas probabilidades iniciais), já que em uma delas (moeda justa) o valor esperado seria zero e nas duas restantes (vieses) ele acredita poder obter vantagem apostando no resultado mais provável, conforme o caso, representando um saldo final positivo.

Finalmente estamos em condições de representar graficamente a situação hipotética em que Wittinho se encontra, conforme abaixo:

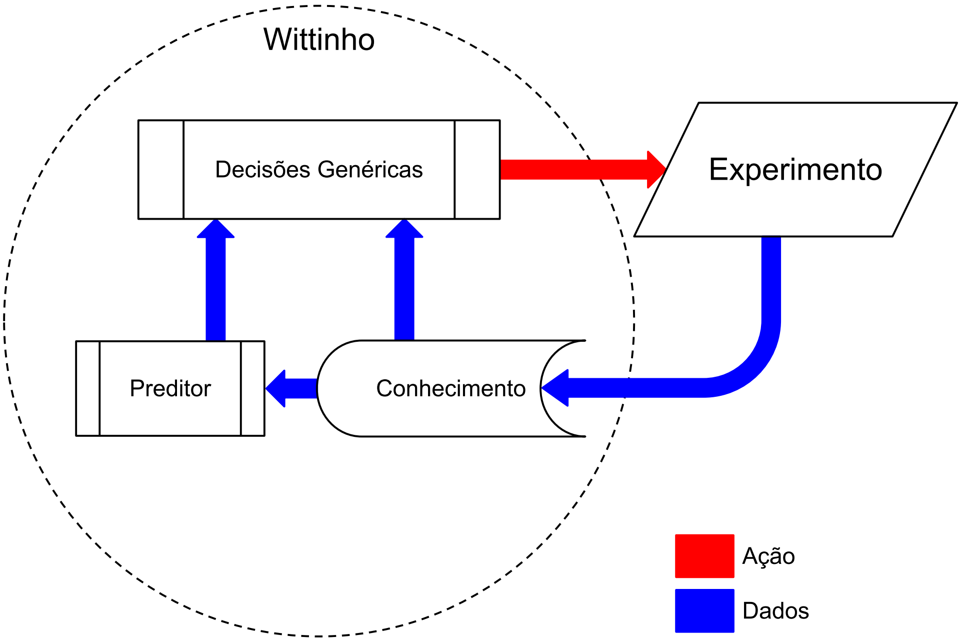

Figura 2.1: Lançamento de moeda e processo de aprendizado.

Na figura, a linha tracejada marca a divisão entre o sujeito cognoscível (Wittinho) e a parte do mundo exterior que corresponde à fonte dos dados relativos à aprendizagem considerada: o experimento.

Internamente ao sujeito estão representados explicitamente o conhecimento e o bloco responsável por realizar as predições, bem como um terceiro objeto responsável pelas decisões genéricas que não são o foco da análise, mas que são necessárias para o funcionamento do processo como um todo.

A partir desse modelo pode-se assumir mais algumas simplificações, sempre observando a manutenção da correspondência entre os principais conceitos abordados no modelo e seus equivalentes no processo de aquisição de conhecimento retratado no experimento. Tal procedimento tem implicações diretas na escolha das ferramentas e conceitos de teoria de conjuntos e teoria de probabilidades a serem usados, conforme se segue.

2.3.3 Teoria de Conjuntos

Na representação em Teoria de Conjuntos há que ser observada uma diferença fundamental entre os blocos que representam ações ou decisões e os que representam passivamente a informação, já que os primeiros se aproximam do que seria equivalente ao processamento de dados a ser usado por Wittinho para tomar as decisões pertinentes e, portanto, serão alvo de mais atenção e considerações.

Os Conjuntos Informativos

Os conjuntos informativos foram concebidos como memórias de armazenamento ou simplesmente fonte da informação (caso do bloco que representa o experimento), esses conjuntos podem ser considerados como constituídos de elementos simples (que não são, eles próprios, conjuntos) ou pares ordenados, caso se deseje indexar as informações em função do tempo ou ordem de amostragem. São eles:

- **Experimento** – embora consista em interação entre o agente e o mundo, já que se trata de lançamentos de moeda, será representado apenas como um conjunto cujos elementos são os resultados possíveis de serem obtidos: cara e coroa.

  Dessa forma, assume-se um nível de conhecimento inferior àquele onde já se disporia da atribuição de probabilidades de obtenção de cada elemento, chegando-se a uma situação equivalente ao caso clássico de sorteio de bolas coloridas de uma urna (com reposição), onde seriam conhecidas que cores há, mas não em que proporções.[2]

$$Experimento = \{Cara, Coroa\}$$

[2] Esse modelo clássico consiste simplesmente em uma urna contendo p bolas pretas e b brancas bem misturadas. Uma bola é extraída aleatoriamente, sua cor é observada e ela é devolvida à urna (ou não, dependendo do problema), sendo o procedimento repetido geralmente um número determinado de vezes. Questões comumente aplicadas ao modelo são: Pode-se inferir a proporção de bolas pretas e brancas a partir de n observações? Com qual grau de confiabilidade? Conhecendo-se p e b, qual a probabilidade de se obter uma sequência específica (por exemplo: uma branca seguida de uma preta)? Se forem observadas apenas n brancas, quão confiante pode-se ficar a respeito da não existência de bolas pretas?

- **Conhecimento** – será representado por um conjunto cujo conteúdo serão os resultados obtidos, em ordem de lançamento. Para tanto, seus elementos serão pares ordenados onde o primeiro elemento identifica o lançamento e o segundo o resultado obtido.

  Dessa forma, iniciaríamos com um conjunto vazio e após o n-ésimo resultado de lançamento (rl_n), teríamos, por exemplo:

$$Conhecimento = \{(1{,}Cara), (2{,}Coroa), (3{,}Coroa), ..., (n{,}Cara)\}$$

ou abreviadamente

$$Conhecimento = \bigcup_{i=1}^{n} \{(i{,}rl_i)\}$$

Talvez cause um pouco de estranhamento a diferença de representação entre esses conjuntos, já que ambos foram classificados como similares.

De fato, o conjunto *experimento* poderia ter sido considerado como uma sequência infinita de resultados ainda não revelados ou, mais em linha com a situação da aposta imaginada, composto por dez pares ordenados de que apenas os índices seriam inicialmente conhecidos.

Isso foi evitado tanto para manter a ideia de escassez de informação, melhor representada pela limitação do conhecimento aos tipos de resultados possíveis, quanto devido ao fato de que os elementos a serem revelados pelos lançamentos já serão devidamente registrados no conjunto *Conhecimento*.

Os Conjuntos Atuadores

Como o próprio nome sugere, a característica peculiar desses conjuntos é que eles influenciam direta ou indiretamente as ações do sujeito frente à situação hipotética como um todo, o que fica claro principalmente no caso do conjunto *Decisões Genéricas*, escolhido para concentrar os detalhes acessórios como a realização dos lançamentos, das apostas, etc.

No caso do *Preditor*, a atuação se limitará ao fornecimento da predição a ser usada, com base em seu conteúdo que, por sua vez, depende do *Conhecimento*. São esses dois conjuntos que representam a aprendizagem aqui considerada, que irá resultar na atitude de Wittinho após cada novo lançamento da moeda, executada pelo *Decisões Genéricas*.

- **Decisões Genéricas** – o conjunto decisões genéricas, como já deve ter ficado claro, foi usado como uma espécie de "coringa" que nos permitiu concentrar nele as funções do sujeito que não são relevantes para o estudo do fenômeno aqui considerado. Justamente por isso é o único bloco representado na figura que não precisa ter detalhado o seu conteúdo, considerando-se que sua atuação será constante durante os dez lançamentos que correspondem ao total de apostas.

  Tal atuação resume-se a, nessa ordem: receber do preditor o próximo resultado em que apostar, fazer a aposta, efetuar o lançamento, observar o resultado e pagar ou receber o valor apostado, conforme o caso.

- **Preditor** – o bloco preditor realiza uma operação de contagem das ocorrências de cara e coroa usando a informação guardada no *Conhecimento*, armazena esse resultado e repassa a ocorrência mais frequente para as *Decisões Genéricas*, mantendo a aposta em Coroa em caso de igualdade.

  Os resultados da contagem serão armazenados como pares ordenados compostos pelo tipo de resultado e pelo respectivo número de ocorrências (Nca = número de caras e Nco = número de coroas), gerando o conjunto:

  $$Preditor = \{(Cara, Nca), (Coroa, Nco)\}$$

  Por exemplo, supondo uma situação onde

  $$Conhecimento = \{(1, Cara), (2, Cara), (3, Coroa), (4, Cara)\}$$

  obteríamos $Preditor = \{(Cara, 3), (Coroa, 1)\}$ e, consequentemente, a predição repassada às decisões genéricas seria “Cara”.

Uma vez que essas representações permitem a construção de um modelo de aquisição de conhecimento indutivo que preserva o problema da indução, conforme será mostrado na análise, resta definir que conceitos de probabilidade serão usados e fazer as considerações necessárias.

2.3.4 Teoria de Probabilidades

> A ciência da lógica atual é familiarizada apenas com coisas certas, impossíveis ou inteiramente duvidosas, nenhuma das quais (felizmente) temos de raciocinar sobre. Portanto, a lógica verdadeira para esse mundo é o cálculo de probabilidades, que leva em conta a magnitude da probabilidade que está, ou deveria estar, na mente o homem razoável.[3] (Maxwell, 1990).

Deixando de lado a diminuição do papel da lógica dedutiva expressada por Maxwell, o caráter contingente presente nas inferências consideradas, que são palco do problema abordado aqui, faz com que seja necessária a verificação tanto da possibilidade de definição e estabelecimento de uma relação entre amostras e predição, quanto da existência de uma base racional equivalente à existente no caso da lógica dedutiva, dando suporte analítico à essa relação.

De posse de um sistema de inferência que atenda a essas características, pode-se partir para a escolha dos conceitos e ferramentas suficientes para tratar o modelo simplificado.

Carnap e Jaynes

Devido ao uso frequente da ideia de probabilidade enquanto tendência a determinado resultado, mesmo antes do desenvolvimento formal realizado por James Bernoulli e Pierre-Simon Laplace, firmou-se uma conexão entre premissas e conclusão que pode ser entendida como sendo de natureza probabilística, ainda que de maneira primitiva

[3] The actual science of logic is conversant at present only with things either certain, impossible, or entirely doubtful, none of which (fortunately) we have to reason on. Therefore the true logic for this world is the calculus of Probabilities, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind.

e não-ótima, sendo possível sua observação mesmo no caso de animais como exemplificado em experimentos com pombos (Skinner, 1948).

Esse uso "instintivo", a que Hume se refere como "hábito", foi dando consistência à ligação entre amostra (premissas) e predição (conclusão), que, se já era forte em termos de senso comum, tornou-se praticamente inquestionável depois da formalização dos cálculos de probabilidades.

Reconhecida a "relação indutiva", a identificação de sua natureza como probabilística implica na possibilidade de assentar sua base sobre os mesmos pilares que dão suporte à teoria de probabilidades, que, sendo aceitos como racionais, justificariam devidamente a conclusão indutiva.

Dessa forma, o estabelecimento dos fundamentos lógicos das probabilidades é a chave para a solução do problema da indução, o que talvez seja o motivo que impossibilitou o próprio Hume de resolvê-lo, já que esses fundamentos foram desenvolvidos bastante defasados em relação às principais regras probabilísticas, sendo desconhecidos à sua época.

Embora a axiomatização desenvolvida por Andrey Nikolaevich Kolmogorov tenha se tornado a referência em Teoria de Probabilidades (Hájek, 2012), a ênfase em lógica usada na abordagem que Rudolf Carnap fez em seu *Logical Foundations of Probability* foi o ponto de partida escolhido para a busca do sistema de inferências indutivas, principalmente devido aos itens expressos logo no primeiro capítulo e tratados como objetivos principais do livro, reproduzidos abaixo:

1. a clarification and, if possible, a definition of the concept of *degree of confirmation*;

2. a clarification of the logical nature of *induction* and, if possible, a construction of a system of *inductive logic*;

3. a clarification of the concept of *probability*.[4]

[4] 1 – a clarificação e, se possível, uma definição do conceito de grau de confirmação; 2 – a clarificação da natureza lógica da indução e, se possível, a construção de um sistema de lógica indutiva; 3 – a clarificação do conceito de probabilidade.

Do ponto de vista dos fundamentos do sistema pretendido, no entanto, a aspiração genérica de Carnap parece, segundo E. T. Jaynes (Jaynes and Bretthorst, 2003, pg. 297), ter sido responsável pelas dificuldades e limitações que ele enfrentou, apesar de ter avançado conceitos importantes como a própria compreensão de probabilidade como grau de confirmação.

Foi o próprio Jaynes, com base em premissas simples assumidas como representando o ideal de racionalidade desejável – os postulados de Pólya-Cox (Arnborg and Sjödin, 2001) – que se propôs a promover a Teoria de Probabilidades à "lógica da ciência", estando em uma posição mais favorável em relação a Carnap quanto ao acesso aos conhecimentos necessários.

Devido a isso, sua abordagem no livro *Probability Theory: the logic of science* foi favorecida em relação à de Carnap, enquanto fonte de onde será extraída a maior parte da informação em Teoria de Probabilidades necessária para a utilização pretendida.

Probabilidade

Conforme visto nas definições expostas na introdução, o sentido do termo probabilidade será restrito ao usado no livro-base de teoria de probabilidades, sustentando-se que os demais usos podem ser devidamente considerados como casos especiais.

Assim, a probabilidade P de um evento E dadas as informações disponíveis I equivale à representação numérica no domínio dos reais do grau de plausibilidade de E ser verdadeira em vista de I. Caso se esteja trabalhando com conjuntos finitos pode-se considerar o domínio como sendo os números racionais, o que é indiferente para a representação usual:

$\mathrm{P(E|I)}$ = probabilidade de E dado I = grau de plausibilidade de E em vista de I

Apesar da diferença de nomenclatura e abordagem geral, essa visão alinha-se à de Carnap quando este define probabilidade do tipo I como o grau de confirmação de uma hipótese com base nas evidências disponíveis.

Desiderata de Pólya-Cox

São critérios de racionalidade aplicados à atribuição de plausibilidades a hipóteses considerando-se as evidências iniciais, bem como à atualização dessas plausibilidades frente a novos dados. Assim, um agente racional, ao descobrir a violação de um deles, teria de revisar o raciocínio e corrigir a atribuição das plausibilidades das inferências relacionadas (Jaynes and Bretthorst, 2003, pg. 9).

A versão apresentada por Jaynes consiste em:

1. **Graus de plausibilidade são representados por números reais**;

2. **Correspondência qualitativa com o senso comum** – sendo A e B duas proposições quaisquer, isso significa que se a plausibilidade de A ser verdadeira aumentar diante da aquisição de dada informação (irrelevante para B), a plausibilidade de $A \wedge B$ também aumentará e a plausibilidade de A ser falsa diminuirá. Adicionalmente, um aumento infinitesimal na plausibilidade de A causará apenas um aumento infinitesimal na plausibilidade das conjunções contendo A, bem como diminuição infinitesimal na plausibilidade da negação de A;

3. **Consistência** – compreendendo as seguintes exigências: a) se pode-se raciocinar de mais de uma maneira sobre algo, deve-se chegar sempre ao mesmo resultado; b) toda informação relevante deve ser levada em conta; c) estados de conhecimento equivalentes devem levar à atribuição das mesmas plausibilidades.

Regra do Produto

Sejam A e B duas proposições cuja plausibilidade depende de uma terceira proposição C. A plausibilidade da conjunção $A \wedge B$, escrita como $P(A \wedge B \mid C)$, relaciona-se com as plausibilidades de A e B, separadamente, da maneira exposta abaixo:

$$P(A \wedge B \mid C) = P(A \mid B \wedge C) \cdot P(B \mid C) = P(B \mid A \wedge C) \cdot P(A \mid C)$$

Regra da Soma

A relação entre a probabilidade de A e sua negação é:

$P(A \mid C) + P(\neg A \mid C) = 1$

Princípio da Indiferença

Sejam $\{H_1, H_2, ..., H_n\}$ hipóteses mutuamente excludentes, exaustivas e indistinguíveis com relação à informação C, suas probabilidades podem ser calculadas como:
$P(H_i|C) = \frac{1}{n}$, onde $1 \leq i \leq n$.

Pode-se derivar, a partir do desiderata 3-c), a necessidade de aplicação do princípio da indiferença na definição das plausibilidades de cada evento, considerando-se, por exemplo, o caso em que eles sejam indistinguíveis com relação a qualquer parâmetro relevante para atribuição de plausibilidades, já que a única regra que se manteria coerente após a troca dos "rótulos" identificadores dos eventos seria a atribuição da mesma plausibilidade a todos eles.

Valor esperado

Em teoria de probabilidades, o valor esperado $E(x)$ para uma variável aleatória discreta x é a média de todos os n valores que ela pode assumir em dado experimento, ponderada pelas probabilidades de obtenção de cada um deles, observando-se que a soma da probabilidades (denominador) é igual a 1. Numa situação prática, pode ser entendido como sendo o valor limite do resultado médio do experimento, quando o número de repetições tende ao infinito. (Hamming, 1991; Ross, 2009)
$E(x) = \sum_{i=1}^{n} [x_i \times P(x_i)]$

Na situação imaginada, supondo 10 apostas fixas de R$ 1, pode-se pensar em um valor esperado *E(a)* para cada uma das hipóteses consideradas, que poderia ser calculado da seguinte maneira:

1. Caso da moeda justa – chance de acerto igual a 1/2 para todas as apostas, levando a: $E(a) = 10 \times \{1 \times \frac{1}{2} - 1 \times \frac{1}{2}\} = 0$;

2. Caso de viés (100%) para cara – chance de acerto igual a 1/2 para a primeira aposta, mas igual a 1 para as 9 restantes, levando a: $E(a) = \{(1 \times \frac{1}{2} - 1 \times \frac{1}{2}) +$

$9 \times (1 \times 1 - 1 \times 0)\} = 9$;

3. Caso de viés (100%) para coroa – chance de acerto igual a 1/2 para a primeira aposta, mas igual a 1 para as 9 restantes, levando a: $E(a) = \{(1 \times \frac{1}{2} - 1 \times \frac{1}{2}) + 9 \times (1 \times 1 - 1 \times 0)\} = 9$;

Finalmente, seguindo o raciocínio inicial de Wittinho quanto à atribuição da probabilidade de 1/3 a cada uma das hipóteses, temos que o "retorno" da decisão geral por participar da aposta compreendendo os 10 lançamentos pode assumir apenas dois valores: R$ 9 (em 2/3 dos casos) e R$ 0 (em 1/3 deles). Isso nos leva ao cálculo do valor esperado geral *E(A)* como:

$$E(A) = \{(0 \times \frac{1}{3} + 9 \times \frac{2}{3})\} = 6$$

Esse cálculo foi checado numericamente por meio de uma simulação simples em computador (o programa consta no apêndice A), consistindo em 1.000.000.000 de repetições (compostas por 10 lançamentos cada). Obteve-se como resultando: Nº caras = 4.999.696.317; Nº coroas = 5.000.303.683 e Nº de acertos = 8.000.192.466. Subtraindo o valor perdido com os erros do ganho com os acertos, tem-se: $1 \times 8.000.192.466 - 1 \times (10.000.000.000 - 8.000.192.466) = 6.000.384.932$, o que dá uma média de aproximadamente R$ 6 por experimento, conforme esperado.

Se considerarmos que Wittinho é neutro com relação à aversão ao risco (na verdade foi suposto inicialmente que ele até estaria disposto a correr algum) podemos considerar *E(A)* como equivalente à *utilidade esperada* dessa ação, que seria a opção escolhida usando o princípio da *maximização da utilidade esperada* (Weirich, 2010), já que a concorrente (não-aposta) tem utilidade igual a zero.

Capítulo 3

Análise e considerações finais

Finalmente estamos em condições de examinar o comportamento do "agente racional" com relação ao processo de aquisição de conhecimento indutivo no universo definido, além de checar a ocorrência de uma instância equivalente ao problema da indução, investigando as possíveis soluções no caso específico, bem como a possibilidade de qualquer generalização.

Observa-se que se está lidando com duas visões sobre o mundo e o processo que nele ocorre: a nossa visão enquanto seres externos e "oniscientes" e a visão do agente (Wittinho).

Espera-se que a racionalidade representada pelos fundamentos do sistema de inferência considerado leve ambas as visões a concordarem em relação às decisões, desde que se desconte a informação extra a que nós temos acesso. Ou seja: ao nos colocarmos "no lugar" do agente, espera-se que cheguemos às mesmas conclusões que ele, já que ambos estamos usando os mesmos critérios de racionalidade.

3.1 A dinâmica do agente

De acordo com a representação usada, o aprendizado sobre que ocorreria o problema da indução é composto pelos seguintes processos a serem repetidos 10 vezes (nessa ordem), conforme estipulado:

1. Predição do resultado do próximo lançamento;

2. Execução do lançamento;

3. Observação do resultado.

É importante lembrar que a predição do primeiro lançamento, como explicado em 2.3.1, será arbitrariamente **Coroa**, uma vez que ainda não há resultados sobre que trabalhar.

Antes de prosseguir ao exame do "comportamento" do agente frente a diferentes resultados do experimento, é necessário dizer que aqui se está considerando a *Definição Geral de Informação* (Floridi, 2011), segundo a qual esta consiste em *dados* + *significado*.

Dessa forma, a participação do preditor no aprendizado estaria justificada quando se considera que ele "extrai" do *Conhecimento* certa informação com base no significado dos dados para o processo conforme estruturado.

Quanto ao papel ativo do preditor, que é informar qual deve ser a próxima aposta, resta verificarmos a equivalência entre a lógica simples usada no preditor do modelo e o raciocínio de Wittinho a respeito das hipóteses e suas implicações, mantendo-se a assunção de viés extremo.

É fácil ver que enquanto for obtido o mesmo resultado do primeiro lançamento, a aposta no resultado mais frequente será equivalente àquela feita sob a suposição de tratar-se de moeda com viés, enquanto qualquer resultado diferente implicaria na conclusão de tratar-se de moeda justa, caso em que a regra usada seria equivalente, em termos de resultado, a apostas aleatórias ou mesmo fixas.

Esse ponto, assim como o aprendizado de maneira geral, ficará mais claro no seguinte exemplo:

Sequência-exemplo: Cara, Cara, Coroa, Cara, Cara.

De acordo com o modelo, antes do primeiro resultado teríamos o seguinte estado relativo ao conhecimento sobre o experimento:

$$Conhecimento = \{\}$$

e

$$Preditor = \{(Cara, 0), (Coroa, 0)\}$$

Com a primeira aposta arbitrada em Coroa, a obtenção de Cara como resultado do primeiro lançamento resultaria em erro na predição e na atualização do estado para:

$$Conhecimento = \{(1, Cara)\}$$

e

$$Preditor = \{(Cara, 1), (Coroa, 0)\}$$

Diante desse novo estado, a regra do maior número de ocorrências leva à predição de Cara para o próximo lançamento, o que é equivalente, em termos de resultado, à assunção de Wittinho da hipótese do respectivo viés como mais provável e à consequente aposta ótima dada a imparcialidade da hipótese de moeda justa.

Assim, seguindo a sequência escolhida, o resultado de Cara para o segundo lançamento consistiria em acerto e no novo estado:

$$Conhecimento = \{(1, Cara), (2, Cara)\}$$

e

$$Preditor = \{(Cara, 2), (Coroa, 0)\}$$

Como Cara continua sendo a ocorrência mais frequente, a aposta para o próximo lançamento seria mantida, resultando em erro frente à obtenção de Coroa (terceira ocorrência conforme a sequência escolhida). Nesse momento ocorreria a eliminação da hipótese segundo a qual trata-se de moeda com viés para Cara, restando apenas a de moeda justa. O novo estado seria atualizado para:

$$Conhecimento = \{(1, Cara), (2, Cara), (3, Coroa)\}$$

e

$$Preditor = \{(Cara, 2), (Coroa, 1)\}$$

Dada a hipótese restante (moeda justa) e a definição de aleatoriedade como incapacidade de predição do resultado do experimento considerado, bastante comum no meio científico (Futuyma, 2005, pg. 225), a regra da ocorrência mais frequente está fadada a obter o mesmo resultado que qualquer outra considerável, por exemplo, como uma que simplesmente apostasse conforme o primeiro lançamento.

Se estivéssemos considerando que a realização da operação de contagem e de mudança de predição resultasse em gasto de energia para o agente, essa regra simplificada seria até mesmo mais eficiente, já que o número de acertos esperados seria o mesmo e o gasto de energia seria menor. Como esse não é o caso, será mantida a regra original, que também pode ser aplicada a casos em que o viés não seja absoluto.

Como há apenas um resultado Coroa na sequência-exemplo e o número de Caras continua maior, as duas apostas seguintes também seriam em Cara e resultariam em mais dois sucessos de predição, levando ao estado final conforme abaixo:

$$Conhecimento = \{(1, Cara), (2, Cara), (3, Coroa), (4, Cara), (5, Cara)\}$$

e

$$Preditor = \{(Cara, 4), (Coroa, 1)\}$$

Como dito anteriormente, observando os dois conjuntos resultantes nota-se que toda a informação já está, de fato, no conjunto *Conhecimento*. Mesmo assim, é o preditor que dá ênfase à característica dos dados considerada relevante para a situação específica imaginada.

Essa característica poderia ser qualquer outra passível de ser "extraída" dos dados obtidos, como os números dos lançamentos que deram Cara (1, 2, 4 e 5), os pares de lançamentos consecutivos com resultados iguais ([1,2] e [4,5]), etc.

Quanto à eficiência das predições, não se pode considerar esse exemplo específico como parâmetro de decisão, já que as assunções do modelo permitem igualmente sequências que resultam em 100% de erros, como mostrado a seguir.

O pior caso

Uma das possibilidades no caso de moeda justa é a sequência: Cara, Coroa, Cara, Coroa, Cara. É fácil ver que a regra do maior número de ocorrências (Coroa em caso de empate) irá gerar uma sequência de predições exatamente oposta à dos resultados, conforme resumido na tabela abaixo:

Tabela 3.1: Sequência pior caso

		Lançamentos				
	Início	1	2	3	4	5
Número de caras (acumulado)	0	1	1	2	2	3
Número de coroas (acumulado)	0	0	1	1	2	2
Predição	xxx	Coroa	Cara	Coroa	Cara	Coroa
Resultado	xxx	Cara	Coroa	Cara	Coroa	Cara

A partir da tabela, observa-se que a predição para o lançamento 1 é obtida do estado inicial (Coroa, em caso de empate). Da mesma forma, para o lançamento 2 os valores considerados são aqueles da coluna 1, levando à escolha de Cara, e assim sucessivamente.

Tal contraste entre os resultados possíveis faz parte da natureza das inferências indutivas entendidas como probabilísticas, no sentido de que mesmo diante de uma "inferência forte" admite-se a possibilidade de falha sem que isso signifique um problema no método como um todo, nem tenha relação direta com o problema da indução que pode ser identificado mesmo nesse modelo simplificado de processo de aprendizado, conforme a seguir.

3.2 Identificando a instância do problema no âmbito do modelo

Relembrando as questões expostas no Explicatum (1.3.3), principalmente quanto as implicações do resultado de um lançamento com relação às hipóteses consideradas e à próxima aposta a ser realizada, observa-se a possibilidade de dúvidas que nos remetem ao problema da indução tão logo seja conhecido o primeiro resultado.

Obviamente não se está falando aqui da eliminação de uma das hipóteses de viés que se segue após a obtenção da primeira Cara ou Coroa, já que ela tem caráter dedutivo, mas da preferência por uma das opções restantes e a efetivação da aposta coerente.

Fazendo referência ao processo aplicado à sequência-exemplo, as perguntas-chave aplicáveis principalmente aos estados 2 e 3 (após o primeiro e segundo lançamento, respectivamente) seriam:

- Se as hipóteses restantes sobre a natureza da moeda são compatíveis com os resultados obtidos, como escolher entre elas? Que tipo de impacto na plausibilidade dessas hipóteses pode haver em face dos resultados específicos obtidos?

- Especificamente sobre a próxima aposta, que parâmetro considerar para escolher entre as opções disponíveis, já que ambas são logicamente possíveis?

Tomando como referência o exemplo clássico do questionamento sobre a continuidade do nascer do sol, sem entrar na questão quantitativa dos números de eventos, podemos considerar a ocorrência de Cara como o nascimento e a de Coroa como o não-nascimento. Da mesma forma, a hipótese de vício seria equiparada à de que o sol continuaria a nascer (regularidade), assim como a de moeda justa equivaleria aquela onde não teríamos nenhuma razão para esperar a continuidade do nascimento.

Sob essa equivalência, as questões postas poderiam ser reescritas como:

- Se tanto a hipótese da regularidade quanto a da não-regularidade do nascer do sol são compatíveis com os resultados obtidos, como escolher entre elas? Que tipo de impacto na plausibilidade dessas hipóteses pode haver dado que o nascimento tem ocorrido até hoje?

- Especificamente sobre o próximo dia, que parâmetro considerar para escolher entre as opções disponíveis (sol nasce $\times$ sol não nasce), já que ambas são logicamente possíveis?

Apesar da ligação óbvia entre a natureza da moeda (ou do sol) e os resultados possíveis (Cara/nascimento e Coroa/não-nascimento), cada um desses entes (hipóteses e resultados) possui seu próprio caráter de plausibilidade, fazendo com que seja possível pensar nas implicações sob o duplo enfoque representado pelas questões.

Diante da equivalência mostrada, pode-se supor que a tentativa de resposta no caso da moeda pode ser posteriormente “traduzida” pelo menos para o caso do nascer do sol, em particular.

3.3 O tratamento do problema

Ao longo do desenvolvimento deste trabalho ocorreram várias frustrações provenientes da falha em seguir de ótimos pontos de partida intuitivos até um desfecho bem-sucedido no tratamento do problema.

Mesmo diante das soluções particularmente sedutoras, como a abordagem de David Stove (Stove, 1986), restava um certo desconforto sobre o que parecia uma divergência implícita entre simpatizantes e críticos que, por nunca vir à tona, criava uma espécie de fosso discursivo onde de um lado uns acham a explicação tão óbvia a ponto de não despertar qualquer interesse em detalhamentos subsequentes, enquanto do outro ela não é convincente a ponto de dar força ao argumento, mas também não parece promissora o suficiente para chamar a devida atenção.

Não foram raras as vezes em que, raciocinando a partir de abordagens distintas, chegava-se a uma só questão que parecia, se não representante do problema da indução em toda a sua generalidade, pelo menos a fonte principal das divergências mencionadas. A pergunta-chave, cujo crédito devo ao grande amigo Allan de Medeiros, é:

- Por que apostar no mais provável?

Sob as constantes investidas dessa questão, começou a tomar forma uma hipótese segundo a qual grande parte da dificuldade do problema da indução consistiria justamente na sua dependência de dois conceitos que são ao mesmo tempo corriqueiros e controversos: a **Racionalidade** e a **Probabilidade**.

A partir dessa hipótese, surge a necessidade de clarear as conexões necessárias à cadeia de raciocínio que visa concluir pela racionalidade de se apostar no mais provável, de forma a forçar a atenção sobre as premissas no caso de o argumento parecer válido, mas não convincente. Assim, espera-se evitar o fenômeno do fosso referido anteriormente através do avanço conforme o roteiro:

1. Definir racionalidade e investigar como ela se relaciona aos processos de decisão em situações de certeza ou incerteza;

2. Checar a racionalidade dos princípios a partir de que pode ser derivado o conceito de probabilidade vista como lógica, assim como de suas implicações;

3. Verificar a adequação da teoria de probabilidades como ferramenta para tomada de decisões e a aplicabilidade ao caso considerado, assim como a consequente racionalidade da solução.

3.3.1 Definindo racionalidade

Devido à importância do conceito, o procedimento prudente a ser adotado em uma definição, ainda que restrita à aplicação aqui pretendida, deveria ser o mesmo usado na identificação do problema da indução: o processo de explicação à moda de Carnap.

Apesar da coerência a ser obtida com essa atitude, a consequência prática equivaleria à suspensão da análise em andamento até que fosse concluída a tarefa de explicação do conceito de Racionalidade, o mesmo sendo repetido com relação ao de Probabilidade, já que segundo a hipótese levantada há pouco, ambos estariam na raiz das dificuldades tradicionalmente encontradas.

Em termos de argumentação, isso equivaleria a transferir o foco da discussão, que é a possibilidade de um argumento em prol da resolução do problema da indução via teoria de probabilidades (o objetivo desse estudo), para as premissas do argumento, que consistem elas próprias em problemas de grande relevância filosófica.

Diante desse impasse, a solução escolhida foi abrir mão de uma representação mais genérica como ponto de partida, por exemplo:

> Racionalidade: em sentido primário, racionalidade é um conceito normativo que os filósofos têm geralmente tentado caracterizar de maneira que, para cada ação, crença ou desejo, se eles são racionais, temos de escolhê-los. Tal caracterização positiva não chegou nem perto de uma aceitação universal porque, frequentemente, várias ações, crenças ou desejos concorrentes contam como racionais.[1] (Audi, 1999, tradução nossa).

Em troca da seguinte referência-chave:

- > Dentre as normas substanciais da razão prática, aquelas da racionalidade instrumental parecem menos controversas para os filósofos. Racionalidade instrumental, em sua forma mais básica, instrui os agentes a agirem conforme os meios que são necessários em relação aos seus

[1]Rationality: In its primary sense, rationality is a normative concept that philosophers have generally tried to characterize in such a way that, for any action, belief, or desire, if it is rational we ought to choose it. No such positive characterization has achieved anything close to universal assent because, often, several competing actions, beliefs, or desires count as rational.

> fins. Na era moderna, essa forma de racionalidade tem sido amplamente vista como o único requerimento isento de problemas para a razão prática.[2] (Wallace, 2009, tradução nossa).

Essa escolha se deu tanto devido à aceitação filosófica quanto por se encaixar perfeitamente ao panorama decisório considerado, principalmente se imaginarmos uma situação ligeiramente diferente para o agente.

Suponhamos que, mantidos todos os outros detalhes, não haja a hipótese de moeda justa. Ora: fica óbvio que a atitude racional, no sentido de ser aquela que levará ao objetivo desejado, será simplesmente apostar na manutenção do primeiro resultado. Nesse caso, conforme mencionado anteriormente, a regra de apostar no evento de maior número de ocorrências será equivalente e também estará em conformidade com os critérios de racionalidade adotados.

A explicação proposta para a perda do sentido de racionalidade ao ser introduzida a incerteza é baseada numa abordagem evolucionário-epistemológica (Bradie and Harms, 2012) com foco tanto sobre as regras responsáveis diretamente pelas escolhas em questão (escolher em qual resultado apostar) – chamemo-las de regras de primeiro nível – quanto sobre a "regra" aplicada para definir o agente quanto à racionalidade: a regra de segundo nível.

Seguindo essa abordagem, a regra racional seria escolhida no nível externo ao agente e, nos casos considerados, equivaleria a apostar no evento que se apresenta em maior proporção dentre as opções disponíveis.

Quando não há incerteza com relação às implicações das escolhas, como no exemplo modificado referido há pouco, há uma coincidência entre as regras de primeiro e segundo nível, já que a certeza pode ser compreendida como razão 1 entre eventos favoráveis e eventos possíveis.

Isso levaria à caracterização como racionais as regras que simplesmente obtiverem sucesso conforme os próprios objetivos a que servem no âmbito do agente. Essa

[2] Among the substantive norms of practical reason, those of instrumental rationality have seemed least controversial to philosophers. Instrumental rationality, in its most basic form, instructs agents to take those means that are necessary in relation to their given ends. In the modern era, this form of rationality has widely been viewed as the single unproblematic requirement of practical reason.

coincidência seria justamente a responsável por ocultar o critério real de racionalidade sob a sombra do caso especial da decisão em situações de certeza, o que dificultaria a identificação do caráter externo sem o qual a compreensão fica prejudicada.

Para tornar mais claro o fenômeno, consideremos uma situação onde a incerteza seja introduzida de forma a evidenciar a diferença entre a abordagem individual e a global, no sentido de que embora não seja possível ter certeza sobre nenhum indivíduo em particular, o resultado geral possa ser conhecido a partir das regras usadas pelos agentes. Imaginemos o seguinte:

Oito pessoas estão diante de uma urna contendo oito bolas. O experimento, a ser repetido 24 vezes, consiste em todos apostarem antecipadamente na cor da bola que esperam retirar da urna e, em seguida, efetuarem a retirada mantendo a bola em seu poder até que todos tenham feito o mesmo. Após a conferência do resultado das apostas, todas as bolas são devolvidas. O seguinte detalhe adicional é informado: os sorteios de número 1 a 8 serão feitos com uma urna que possui 5 bolas pretas e 3 brancas (urna **5/3**), os de número 9 a 16 com uma que possui 2 pretas e 6 brancas (urna **2/6**) e os de número 17 a 24 com uma contendo 7 pretas e 1 branca (urna **7/1**).

Considere-se que as pessoas sabem as quantidades de bolas pretas e brancas em todos os momentos do experimento, mas não têm nenhuma noção de qualquer conexão entre as proporções e a melhor aposta, além de não terem nenhuma experiência anterior capaz de sugerir essa ligação. Vamos partir da assunção de que elas simplesmente irão se comportar conforme um determinado "hábito" no decorrer do experimento, que consistirá basicamente em apostar na repetição do último resultado.

Se não há, segundo Hume, justificação racional para qualquer hábito no contexto de decisões sob incerteza, o nosso é, a priori, tão bom quanto qualquer outro. Esse comportamento também pode ser entendido como a tradicional assunção de que o passado (digamos que as pessoas têm um grave problema de memória) é similar ao futuro.

Para iniciar a sequência de maneira equilibrada, as pessoas (P1, P2, ..., P8) serão divididas em dois grupos da seguinte maneira: P1 a P4 apostarão na obtenção de bola preta, enquanto P5 a P8 apostarão na obtenção da branca. Após os primeiros resultados, o hábito descrito entrará em ação, gerando a seguinte tabela:

Tabela 3.2: Tabela de evolução de escolhas confome hábito

	Pessoas																
	P1		**P2**		**P3**		**P4**		**P5**		**P6**		**P7**		**P8**		
S	**A**	**R**	**A**	**R**	**A**	**R**	**A**	**R**	**A**	**R**	**A**	**R**	**A**	**R**	**A**	**R**	**P/B**
1	P	P	P	B	P	P	P	B	B	P	B	B	B	P	B	P	5/3
2	P	B	B	P	P	P	B	P	P	B	B	P	P	P	P	B	
3	B	P	P	P	P	B	P	B	B	P	P	P	P	B	B	P	
4	P	P	P	P	B	P	B	P	P	B	P	B	B	B	P	P	
5	P	B	P	P	P	P	P	B	B	P	B	P	B	P	P	B	
6	B	P	P	B	P	B	B	P	P	P	P	B	P	P	B	P	
7	P	B	B	P	B	P	P	P	P	B	B	B	P	P	P	P	
8	B	B	P	P	P	P	P	P	B	B	B	P	P	P	P	B	
9	B	B	P	P	P	B	P	B	B	P	P	B	P	B	B	B	2/6
10	B	B	P	P	B	B	B	P	P	B	B	B	B	B	B	B	
11	B	B	P	B	B	B	P	B	B	P	B	P	B	B	B	B	
12	B	P	B	B	B	B	B	B	P	B	P	B	B	B	B	P	
13	P	B	B	P	B	B	B	B	B	B	B	B	B	P	P	B	
14	B	B	P	B	B	P	B	B	B	B	B	P	P	B	B	B	
15	B	B	B	B	P	P	B	B	B	B	P	B	B	P	B	B	
16	B	B	B	B	P	B	B	P	B	B	B	B	P	B	B	P	
17	B	P	B	P	B	P	P	P	B	P	B	P	B	P	P	B	7/1
18	P	P	P	B	P	P	P	P	P	P	P	P	P	P	B	P	
19	P	P	B	P	P	P	P	P	P	P	P	P	P	B	P	P	
20	P	P	P	P	P	P	P	P	P	P	P	B	B	P	P	P	
21	P	B	P	P	P	P	P	P	P	P	B	P	P	P	P	P	
22	B	P	P	P	P	P	P	P	P	B	P	P	P	P	P	P	
23	P	B	P	P	P	P	P	P	B	P	P	P	P	P	P	P	
24	B	P	P	P	P	P	P	B	P	P	P	P	P	P	P	P	

S = sorteio; A = aposta; R = resultado (erros em vermelho)

Iniciemos as observações do nosso ponto de vista privilegiado, de modo a apontar algumas características globais importantes:

É notável a convergência do hábito para o ato de apostar no evento mais provável à medida que as proporções das bolas se afastam de 50%, como pode-se observar comparando sucessivamente os casos 5/3, 2/6 e 7/1, mesmo sob restrição tão severa de memória (apenas o último resultado).

Tal resultado deve-se ao fato necessário de que em qualquer sorteio há mais maneiras de se acertar apostando no evento em maior proporção, o inverso ocorrendo com as maneiras de errar.

Tomando o caso 7/1, em que isso fica mais óbvio, se determinada pessoa faz a aposta na bola branca, é fácil ver que há 8 maneiras de distribuí-la, das quais apenas uma resultará em acerto (aquela em que ela efetivamente recebê-la).

É essa assimetria entre maneiras de errar e de acertar que faz com que as apostas tenham uma tendência a fixarem-se nas bolas pretas, mesmo tendo iniciado as sequências relativas a essa urna com 7 apostas nas brancas (Sorteio 17), devido ao viés inicial fornecido pela urna anterior (2/6).

Consideremos agora a visão dos indivíduos sobre a racionalidade de suas próprias apostas, ainda supondo que eles não foram capazes de estabelecer uma relação entre as proporções de bolas na urna (no caso, a 7/1) e o resultado geral esperado em decorrência dela.

Se as pessoas apenas contarem com suas próprias experiências e não houver um caráter público do conhecimento (não havendo troca de informação entre os agentes), o critério racional aplicado por cada agente individualmente permite a construção de um "quadro de racionalidades" em que cada resultado de sorteio fará com que 7 pessoas considerem a aposta em preta como racional e apenas uma no caso da branca, estando elas justificadas se levarmos em conta a limitação de informação.

Se relaxarmos a restrição quanto à comunicação e considerarmos que os agentes que errarem as apostas possam perguntar aos demais em que devem apostar, veremos imediatamente que as sugestões refletirão de maneira indireta as proporções de bolas na urna, de modo que, mesmo sem apelar para a decisão via maioria, a escolha aleatória de um conselheiro irá levar, na maior parte das vezes, à aposta nas bolas pretas.

Por outro lado, se imaginarmos que a escolha da dica possa ser realizada conforme a sua popularidade, já estaremos indiretamente caminhando em direção à aposta no evento em maior proporção, o que explicaria, pelo menos nessa situação específica, a conexão entre publicidade e racionalidade normalmente estabelecida.

De maneira geral, ainda que admitamos a possibilidade de procedimentos racionais diferentes, um ambiente que atribuir a continuidade ou a replicação de uma regra de inferência[3] ao sucesso de suas predições, exerce uma pressão que força a população de regras a convergir para apostar no evento em maior proporção, quando ele existe, o que faz sentido se pensarmos que a diferença em proporções consiste ela mesma em informação que pode ser levada em conta.

3.3.2 A racionalidade da Teoria de Probabilidades

Uma vez que a regra racional geral consiste em apostar conforme a maior proporção, da qual a decisão em situações de certeza (proporção de 100%) é um caso especial, pode-se avançar ao nível de incerteza mais próximo daquele característico do processo de aquisição de conhecimento científico e tentar entender como é possível manter a racionalidade das decisões.

Os casos reais geralmente permitem apenas hipóteses sobre as proporções dos eventos possíveis, como na situação hipotética enfrentada pelo personagem Wittinho. No exemplo da urna, isso equivaleria a informar apenas a quantidade total e cores das bolas, obrigando o agente a ter de considerar inicialmente todas as possibilidades de proporções (8 pretas, 7 pretas e 1 branca, 6 pretas e duas brancas, ..., 1 preta e 7 brancas, 8 brancas).

Esse novo ponto de partida levanta dois desafios para qualquer método que objetive se manter fiel ao raciocínio anterior, onde as proporções não só eram conhecidas como atuavam como critério racional de decisão. Ilustremos esses desafios por meio de uma comparação direta entre as situações, com relação aos seguintes pontos:

- **A primeira aposta** – enquanto na primeira situação há apenas uma aposta ótima (racional), no segundo todas as hipóteses de proporções possíveis têm um equivalente simétrico a partir do qual a escolha oposta poderia ser inferida. Assim sendo, é desejável que haja algum meio que permita a escolha inicial

[3] Pensando numa população onde os indivíduos são regras como, por exemplo, as que geram as apostas nas cores das bolas.

dentre essas possibilidades, sendo o próprio critério de apostar na hipótese em maior proporção o candidato principal a ponto de partida.

- **Resultados** $\times$ **Hipóteses** – se considerarmos o processo de retirada como não favorecendo nenhuma das 8 bolas, ou seja: se ele for "aleatório", o fato de o agente retirar n bolas brancas consecutivas no caso 7/1 não faz com que a atitude racional deixe de ser apostar nas pretas desde que as premissas estejam corretas (escolha aleatória e proporção 7/1), o que é comumente caracterizado como independência entre eventos.

  Ao abrirmos mão da certeza sobre o conteúdo da urna, entretanto, temos de levar em conta o grau de suporte que os resultados obtidos dão a cada uma das hipóteses, desde os casos óbvios em que de uma urna contendo apenas bolas brancas não se pode extrair uma preta, até os menos óbvios em que se deve esperar que uma urna similar à do caso 7/1 origine mais sorteios de bolas pretas que uma cujo conteúdo é oposto (1 preta e 7 brancas, ou 1/7).

  Esse grau de evidência que determinado resultado fornece para cada hipótese acaba por eliminar o caráter de independência entre eventos, observando que essa "independência lógica", mesmo no caso de certeza quanto às proporções, não deve ser confundida com uma propriedade real do conjunto [urna + processo de sorteio], mas apenas como resultado da ausência de qualquer informação a respeito.

Mesmo após os desenvolvimentos mais recentes em Teoria de Probabilidades, o primeiro ponto levantado – o problema dos *priors* – ainda é considerado como um campo de estudo que requer muito desenvolvimento (Jaynes and Bretthorst, 2003, pg. 88). É também a maior fonte de críticas ao uso da Teoria de Probabilidades enquanto lógica, independentemente de sua aplicação ao problema da indução (Franklin, 2001, pg. 278).

A despeito disso, há casos em que a aplicação do princípio da indiferença se mostra apropriada e, além disso, sua conclusão (atribuição de mesma probabilidade aos eventos) também tem sido vista como um caso particular da aplicação do princípio da máxima entropia (Jaynes and Bretthorst, 2003, pg. 353) aos casos discretos onde a informação disponível é perfeitamente simétrica com relação às possibilidades.

Nesses casos, o princípio da máxima entropia pode ser entendido como informando que a distribuição uniforme é a mais numerosa ao se considerar o espaço de distribuições possíveis e, portanto, aquela que mais se espera encontrar.

Os casos considerados, desde o exemplo prático da moeda aos exemplos auxiliares das urnas, fazem parte da classe de experimentos conhecida como ensaios de Bernoulli. Como essa classe permite a aplicação dos princípios referidos anteriormente, é possível escolher uma proporção específica de moedas ou bolas (conforme o caso) como hipótese inicial.

O próximo passo para restaurar a posse do critério de racionalidade é relativo ao segundo problema mencionado. Consiste em considerar nos cálculos a influência que a obtenção de novos dados tem sobre a plausibilidade da hipótese inicial e, consequentemente, sobre a dos resultados posteriores. Em outras palavras, corrigir as crenças nas probabilidades (ou proporções) de acordo com a aquisição de informação adicional representada pelo conhecimento dos resultados dos sorteios (ou lançamentos, no caso de moedas).

Embora as clássicas regras do produto e da soma mostradas em 2.3.4 já venham sendo usadas para essa finalidade há bastante tempo, a possibilidade de derivá-las a partir dos *Desiderata de Pólya-Cox* (Jaynes and Bretthorst, 2003, pg. 24-34) fornece todo um contexto de racionalidade particularmente relevante em nossa abordagem. Além da "validação racional" das regras, uma consequência exprimida particularmente pelo item 3-c) dos desiderata (ver 2.3.4) é que quaisquer regras alternativas para atribuição de probabilidades que partam do mesmo estado cognoscível terão necessariamente de chegar às mesmas conclusões!

Ao resolver ambos os problemas levantados, a Teoria de Probabilidades entendida em sentido amplo (incluindo os princípios usados para definir as proporções iniciais) é capaz de estender o raciocínio que parte da certeza quanto à constituição do conjunto de elementos a casos mais genéricos e que representam melhor os processos de inferência sob incerteza, tais como os exemplos considerados nesse trabalho.

3.3.3 Análise quantitativa

Após a "validação racional" das ferramentas escolhidas, pode-se finalmente aplicá-las ao modelo de aprendizado considerado, de forma a observar alguns indicadores quantitativos e sua conexão com as decisões relacionadas.

O conhecimento inicial

Embora o papel da informação, nem sempre explícita na formulação do problema e de que se derivam as probabilidades iniciais, já tenha sido comentado em 2.3.1, a possibilidade de representação numérica permite uma visão melhor das assunções muitas vezes implícitas que estão por trás das primeiras atribuições de probabilidades às hipóteses e das quais depende toda a cadeia de raciocínio indutivo.

Comparemos os seguintes casos a respeito das possíveis origens de uma moeda a ser usada para realizar lançamentos:

1. Uma caixa contém 3 moedas, sendo uma delas comum (possui uma face cara e uma coroa), uma com duas caras e uma com duas coroas;

2. Uma moeda estilizada será fabricada a partir de um disco de dimensões apropriadas, recebendo uma marcação de cada um de seus lados. Para cada lado, a marcação consistirá em um desenho de um rosto ou um de uma coroa e será feita independentemente do que for marcado no outro.

3. Uma pessoa resolve trocar, numa loja, uma cédula de R$ 2 por duas moedas de R$ 1, de modo a usar uma delas para fazer lançamentos.

O primeiro item é um equivalente, em termos de informação, ao estado de conhecimento escolhido arbitrariamente para Wittinho. Assim, também resulta na atribuição de probabilidades de $1/3$ para casa hipótese via princípio da indiferença.

Já no segundo caso, apesar do enunciado mais longo, a situação é colocada de forma a dar ênfase apenas às diferentes maneiras de fabricação, sem beneficiar nenhuma delas em detrimento das outras, aproximando-se da abordagem via estados possíveis de Carnap e, consequentemente, assumindo menos informações sobre o mundo. Seja a representação de possibilidades dada pela tabela abaixo:

Tabela 3.3: Possibilidades de fabricação da moeda, conforme informações fornecidas

Possibilidades	**Lado 1**	**Lado 2**	**Moedas resultantes**
a	Cara	Cara	Viés p/ cara
b	Cara	Coroa	Justa
c	Coroa	Cara	Justa
d	Coroa	Coroa	Viés p/ coroa

Observa-se que os itens b e c são equivalentes a moedas justas, o que faz com que essa característica esteja presente em metade das possibilidades. Apesar de mais próximo das nossas expectativas de senso comum, já que normalmente acreditamos mais na honestidade da moeda que em tratar-se de uma com viés, é o terceiro caso que pode ser considerado como fonte de uma distribuição de probabilidades iniciais mais próxima do que se espera verificar na maioria das pessoas.

Em termos comparativos isso significa que as probabilidades iniciais geradas pelo caso 2 fariam com que o agente racional cujas crenças fossem representadas por ele passasse a acreditar no viés da moeda muito mais facilmente que um cidadão comum, já que bastaria que os dois primeiros lançamentos dessem o mesmo resultado, como poderá ser visto nos cálculos abaixo.

Dadas as informações:

- X = caso 2 (manufatura da moeda estilizada).
- $P(Justa \mid Cara \wedge X)$ – probabilidade de moeda ser justa, dada a obtenção de cara e a informação inicial;
- $P(Cara \mid Justa \wedge X)$ – probabilidade de obtenção de cara, dado que a moeda é justa e considerando-se a informação inicial;

- $P(Justa \mid X)$ – probabilidade de se obter uma moeda justa dada apenas a informação inicial;

- $P(Cara \mid X)$ – probabilidade de obtenção de cara dada apenas a informação inicial;

Nesse caso, pode-se usar a regra do produto na forma:

$$P(Justa \mid Cara \wedge X) = \frac{P(Cara \mid Justa \wedge X) \times P(Justa \mid X)}{P(Cara \mid X)}$$

Cuja substituição de valores leva a:

$$P(Justa \mid Cara \wedge X) = \frac{\frac{1}{2} \times \frac{1}{2}}{\left(1 \times \frac{1}{4}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) + \left(0 \times \frac{1}{4}\right)} = \frac{1}{2}$$

Da mesma forma, para as moedas tipo a e d, tem-se:

$$P(a \mid Cara \wedge X) = \frac{1 \times \frac{1}{4}}{\left(1 \times \frac{1}{4}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) + \left(0 \times \frac{1}{4}\right)} = \frac{1}{2}$$

e

$$P(d \mid Cara \wedge X) = \frac{0 \times \frac{1}{4}}{\left(1 \times \frac{1}{4}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) + \left(0 \times \frac{1}{4}\right)} = 0$$

Como já foi eliminada (atribuição de probabilidade 0) a hipótese representada pelo item d na tabela, basta refazermos os cálculos de modo a contabilizar a influência do segundo resultado, mas agora considerando as probabilidades atualizadas de a moeda ser do tipo a ou *justa* como sendo $1/2$, de forma a obter:

$$P(a \mid Cara \wedge Cara \wedge X) = \frac{1 \times \frac{1}{2}}{\left(1 \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) + (0 \times 0)} = \frac{2}{3}$$

e

$$P(Justa \mid Cara \wedge Cara \wedge X) = \frac{\frac{1}{2} \times \frac{1}{2}}{\left(1 \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) + (0 \times 0)} = \frac{1}{3}$$

Se continuarmos admitindo as assunções sobre a racionalidade de apostar no evento em maior proporção, bem como sobre as maneiras de manipulação dessas proporções frente à obtenção de novos dados, torna-se clara a mudança de atitude do agente cujo conhecimento inicial é representado pela tabela 3.3 quanto à hipótese sobre a natureza da moeda obtida, devendo ele passar a apostar considerando que a moeda tem viés para cara.

Seguindo o mesmo raciocínio, o fato de uma pessoa hipotética não mudar de opinião a respeito do tipo da moeda após 4 caras seguidas poderia ser entendido como irracional apenas se supusermos que ela tem exatamente as mesmas informações representadas pela tabela, o que não é o caso, por exemplo, na terceira situação imaginada (obtenção de moeda em loja).

Apesar de não ter sido o ponto principal nesse estudo, a atribuição da probabilidades iniciais a serem consideradas nos cálculos parece evidenciar uma relação entre a crença nas proporções entre as hipóteses e o "grau de crença" (*degree of belief*) em cada uma delas, pelo menos do ponto de vista funcional, já que a atribuição de probabilidade baixa a uma hipótese requer mais resultados favoráveis a ela para que o agente passe a adotá-la (quando sua probabilidade superar a das concorrentes).

Evidentemente há casos em que a atribuição ideal de probabilidades é facilmente identificada como correspondendo às proporções reais dos eventos, como aqueles em que o conteúdo das urnas é conhecido, entretanto, não é necessariamente a correção dessa atribuição que definirá o caráter racional do agente, da mesma forma que um argumento dedutivamente válido não requer a verdade das premissas.

Por último, uma vez supondo a representação da informação inicial através das probabilidades iniciais, quaisquer estados cognoscíveis que resultarem em distribuições equivalentes (a despeito da nomenclatura) podem ser entendidos como sendo equivalentes para os fins dos cálculos probabilísticos, conforme esperado devido ao desiderata *3c* em 2.3.4.

Foi essa propriedade a responsável pela afirmação feita há pouco de que o caso 1 na lista das possíveis origens da moeda seria equivalente ao conhecimento de Wittinho a respeito das probabilidades relativas ao seu caso específico, mesmo tendo sido expresso anteriormente apenas como uma sensação de indecisão a respeito de qualquer uma das alternativas. Passemos agora às demais considerações qualitativas.

Os casos 1 e 2 revisitados

O estado de conhecimento inicial de Wittinho, conforme observado anteriormente, pode ser representado pela tabela abaixo:

Tabela 3.4: Probabilidades iniciais das hipóteses

	Hipóteses		
	1 (100% *ca*)	2 (50% *ca* e 50% *co*)	3 (100% *co*)
$P(H_n \mid X)$	$1/3$	$1/3$	$1/3$
$P(ca \mid H_n)$	1	$1/2$	0
$P(co \mid H_n)$	0	$1/2$	1

H_n=Hipótese n; X=constituição da urna; ca=cara; co=coroa.

Onde:

- $P(H_n \mid X)$ – probabilidade da hipótese n ser verdadeira, sabendo-se que a moeda correspondente é obtida de uma caixa contendo exatamente uma justa, uma com duas caras e uma com duas coroas, por meio de sorteio aleatório;
- $P(ca \mid H_n)$ – probabilidade de obtenção de cara a partir da moeda correspondente à hipótese H_n;
- $P(co \mid H_n)$ – probabilidade de obtenção de coroa a partir da moeda correspondente à hipótese H_n;

Considerando-se R_i como o i-ésimo resultado (cara ou coroa) e lembrando que X consiste na informação inicial sobre a urna, pode-se aplicar novamente a regra do produto para atualizar as probabilidades de cada hipótese frente à obtenção de R_i conforme abaixo:

$$P(H_n \mid R_i \wedge R_{i-1} \wedge ... \wedge R_1 \wedge X) = \frac{P(R_i \mid H_n \wedge R_{i-1} \wedge ... \wedge R_1 \wedge X) \times P(H_n \mid R_{i-1} \wedge ... \wedge R_1 \wedge X)}{P(R_i \mid R_{i-1} \wedge ... \wedge R_1 \wedge X)}$$

O primeiro caso (2.3.1) aborda de maneira simples como a obtenção de cara no primeiro lançamento poderia ser usada para decidir pela próxima aposta, concluindo que se deve apostar na repetição desse resultado com uma confiança crescente enquanto não ocorrer nenhuma coroa.

Supondo que os 9 lançamentos seguintes acabassem por também resultar em cara, a aplicação sucessiva da regra do produto conforme mostrada acima geraria os dados mostrados na seguinte tabela:

Tabela 3.5: Atualização das probabilidades das hipóteses no caso 1

Resultados		$P(H_1 \mid R_i \sim R_1 \wedge X)$	$P(H_2 \mid R_i \sim R_1 \wedge X)$	$P(H_3 \mid R_i \sim R_1 \wedge X)$
1	ca	0,66667	0,33333	0,00000
2	ca	0,80000	0,20000	0,00000
3	ca	0,88889	0,11111	0,00000
4	ca	0,94118	0,05882	0,00000
5	ca	0,96970	0,03030	0,00000
6	ca	0,98462	0,01538	0,00000
7	ca	0,99225	0,00775	0,00000
8	ca	0,99611	0,00389	0,00000
9	ca	0,99805	0,00195	0,00000
10	ca	0,99902	0,00098	0,00000

As probabilidades atualizadas após o primeiro resultado (linha 1) equivalem justamente ao raciocínio ilustrado anteriormente por meio da representação do conjunto de lançamentos que resultaram em cara, cuja composição tenderia a consistir em 2/3 (aproximadamente 0,66667) de moedas viciadas e 1/3 (aproximadamente 0,33333) de moedas justas.

Graças à facilidade dos cálculos probabilísticos e à possibilidade de estabelecer uma correspondência quantitativa entre a crença na probabilidade de determinada hipótese e o grau de crença em sua veracidade, pode-se entender o processo representado pela tabela como justamente a criação e fortalecimento do "hábito" correspondente a esperar a continuidade da obtenção de cara como resultado.

O caso 2 (2.3.1) é praticamente idêntico ao 1 em termos de cálculo, a não ser pelo fato de que os nove primeiros resultados consistem em coroa, embora o último seja cara. A partir desses dados, a tabela resultante é:

Tabela 3.6: Atualização das probabilidades das hipóteses no caso 2

Resultados		$P(H_1 \mid R_i \sim R_1 \wedge X)$	$P(H_2 \mid R_i \sim R_1 \wedge X)$	$P(H_3 \mid R_i \sim R_1 \wedge X)$
1	co	0,00000	0,33333	0,66667
2	co	0,00000	0,20000	0,80000
3	co	0,00000	0,11111	0,88889
4	co	0,00000	0,05882	0,94118
5	co	0,00000	0,03030	0,96970
6	co	0,00000	0,01538	0,98462
7	co	0,00000	0,00775	0,99225
8	co	0,00000	0,00389	0,99611
9	co	0,00000	0,00195	0,99805
10	ca	0,00000	1,00000	0,00000

A partir da tabela acima, pode-se observar que as regras empregadas também são capazes de gerar inferências necessárias, desde que alimentadas com os dados suficientes para tanto. Esse fenômeno está representado no caso da atribuição de plausibilidade 0 para a hipótese de tratar-se de moeda com viés de 100% para cara (ou para coroa, no caso 1), bem como na atribuição de plausibilidade 1 para a hipótese de moeda justa ao mesmo tempo em que a hipótese de viés para coroa também recebe plausibilidade 0, após a obtenção do resultado do lançamento 10.

Do ponto de vista da coerência, não há resultados posteriores capazes de mudar o estado de certeza a respeito de alguma hipótese, a não ser questionando as informações usadas para se chegar a ele, quer sejam as probabilidades iniciais ou algum resultado de lançamento.

Isso está de acordo com o que se esperaria no caso de um argumento dedutivamente válido e retrata o papel de extensão da lógica clássica exercido pela teoria de probabilidades conforme promovido por Jaynes (Jaynes and Bretthorst, 2003, pg. 12).

Para fechar as considerações sobre o caso 2, abordemos as seguintes questões deixadas em aberto anteriormente:

- Qual a razoabilidade de se obter predições acertadas usando hipóteses falsas?
- O que está por trás da dificuldade intuitiva em lidar com eventos raros e de que forma a teoria de probabilidades pode ajudar?

De modo a manter a coerência e evitar extrapolação de competência, as respostas às questões estarão obviamente limitadas ao alcance do critério-base de racionalidade adotado (aposta no evento em maior proporção – conforme 3.3.1), bem como às conclusões possíveis a partir do uso das ferramentas probabilísticas de atualização das hipóteses em face da obtenção de dados.

Se raciocinarmos partindo da probabilidade calculada e representarmos o problema como sendo equivalente à escolha por pertencer a um entre dois conjuntos distintos, fica claro que o conjunto dos agentes que seguirem os critérios racionais e que acabarem por estar errados será muito menor que o daqueles que não os seguirem e acabarem por estar certos.

A primeira questão colocada é justamente um caso especial em que os critérios racionais foram seguidos, mas levaram à crença na hipótese errada (tratar-se de uma moeda com viés). Dada a natureza auto-corretiva do sistema, é de se esperar que a manutenção de uma hipótese, seja qual for, requer dados que a reforcem ou pelo menos que sejam compatíveis com ela, fazendo com que a permanência no erro sobre a natureza da moeda implique paradoxalmente no acerto das previsões.

Isso não muda o fato de tratar-se de um caso que torna-se mais raro à medida que mais informações são obtidas e que a informação de que dispomos sobre o processo apenas nos permite escolher entre as estratégias de ação, mas não entre um resultado específico dentre as possibilidades.

A resposta à segunda pergunta já foi sugerida parcialmente em 1.1 ao considerarmos a questão da sobrevivência como limite mínimo a ser obedecido. Se pensarmos apenas na capacidade reprodutiva do agente sob a influência das decisões relativas à ocorrência de um evento, fica fácil ver que ele se torna menos relevante conforme a sua raridade.

Além disso, do ponto de vista de aquisição e processamento de informação, eventos raros também exigem um nível de conhecimento muitas vezes fora do alcance da intuição do homem tribal, desenvolvida em um ambiente menos complexo e sob a influência de menos agentes. Com a grande disponibilidade de informação alcançada nas últimas décadas, a teoria de probabilidades torna possível a tomada de decisões capazes de aproveitar a grande escala das sociedades atuais somando pequenas possibilidades de ganho e eliminando pequenas perdas difíceis de lidar intuitivamente.

Por fim, no que diz respeito à possibilidade de comparação entre crença nas proporções e grau de crença, agora dispomos de ferramentas capazes de representar quantitativamente crenças diferentes, além de julgá-las quanto à racionalidade e de fornecer um ponto de partida para estudos sobre suas origens e sobre a maneira de tornar a intuição compatível com o novo ambiente de informação no qual estamos inseridos.

A sequência-exemplo e o pior caso

Antes de fazer os cálculos relativos ao agente modelado, faz-se necessário ressaltar uma diferença importante que foi mencionada apenas de maneira rápida anteriormente.

Como afirmado em 3.1, o desempenho do agente atuante conforme a regra de apostar no evento de maior ocorrência será equivalente ao comportamento de Wittinho. Apesar disso, fica claro que há uma "disposição" de mudar a aposta sempre que o número de caras tornar-se superior (ou deixar de sê-lo) ao de coroa, enquanto Wittinho, diante das hipóteses consideradas como possíveis, não teria preferência por nenhuma aposta tão logo fossem obtidos resultados diferentes, o que implicariam se estar diante de moeda justa.

Traduzindo em termos de crença inicial, tal susceptibilidade em mudar de aposta conforme a mínima diferença em favor de um dos eventos indica que o agente atribui uma probabilidade inicial igual a zero para a hipótese de moeda justa.

Do ponto de vista lógico, entretanto, devemos fazer outra alteração antes de poder representar a dinâmica usando a tabela de cálculos, conforme vem sendo feito. Ao eliminarmos a possibilidade de moeda justa, estamos automaticamente limitando os resultados dos lançamentos a sequências consistindo de um mesmo resultado, incapazes de representar tanto a sequência-exemplo quanto o pior caso considerado.

Dessa forma, para permitir a obtenção de resultados distintos em uma mesma sequência, é necessário que eliminemos o caráter extremo do viés, o que faremos de forma a manter a simetria quanto às possibilidades restantes, ou seja: se supusermos que um dos vieses é de 99 caras em 100 lançamentos, o outro será inversamente de 99 coroas em 100 lançamentos.

Finalmente, após essas observações, passemos à tabela a seguir, que reflete duas possibilidades distintas de vieses:

Tabela 3.7: Inferências relativas à sequência-exemplo

		Viés forte		**Viés fraco**	
		H_1: viés p/ cara	H_2: viés p/ coroa	H_1: viés p/ cara	H_2: viés p/ coroa
$P(H_n \mid X)$		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$P(ca \mid H_n)$		$\frac{999}{1000}$	$\frac{1}{1000}$	$\frac{501}{1000}$	$\frac{499}{1000}$
$P(co \mid H_n)$		$\frac{1}{1000}$	$\frac{999}{1000}$	$\frac{499}{1000}$	$\frac{501}{1000}$
Resultados		$P(H_1 \mid R \wedge X)$	$P(H_2 \mid R \wedge X)$	$P(H_1 \mid R \wedge X)$	$P(H_2 \mid R \wedge X)$
1	*ca*	0,999000000	0,001000000	0,501000000	0,499000000
2	*ca*	0,999998998	0,000001002	0,501999992	0,498000008
3	*co*	0,999000000	0,001000000	0,501000000	0,499000000
4	*ca*	0,999998998	0,000001002	0,501999992	0,498000008
5	*ca*	0,999999999	0,000000001	0,502999968	0,497000032

Considerando a probabilidade como razão de ocorrência de um evento relativo ao número de repetições de determinado experimento, podemos notar que após qualquer resultado de lançamento a única opção sensata continuou sendo a aposta em cara, o que implica na racionalidade de apostar conforme a maior proporção da amostra, se considerarmos essa sequência específica e as premissas adotadas.

É interessante observar que a magnitude do viés não interfere na escolha, já que não há discordância sobre que resultado favorece qual hipótese, apesar de o grau de confirmação ser distinto quando comparamos os casos de viés forte e fraco. Como não se pode optar por não apostar nem se pode mudar o valor da aposta, essa diferença acaba por não gerar nenhum resultado prático no âmbito da situação imaginada.

Caso não houvessem essas restrições, entretanto, seria necessário recorrer novamente ao cálculo do valor esperado de modo a definir a partir de que valor[4], a cada estado cognoscível, o agente estaria disposto a apostar.

Diante dessa nova situação seria possível observar uma atitude diferente relativa à crença em um viés forte, onde o agente estaria disposto a um maior risco a cada aposta.

Quanto ao pior caso passível de ser obtido a partir das regras escolhidas, se usarmos as mesmas premissas adaptadas obtemos a tabela seguinte:

Tabela 3.8: Inferências relativas ao pior caso

		Viés forte		**Viés fraco**	
		H_1: viés p/ cara	H_2: viés p/ coroa	H_1: viés p/ cara	H_2: viés p/ coroa
$P(H_n \mid X)$		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$P(ca \mid H_n)$		$\frac{999}{1000}$	$\frac{1}{1000}$	$\frac{501}{1000}$	$\frac{499}{1000}$
$P(co \mid H_n)$		$\frac{1}{1000}$	$\frac{999}{1000}$	$\frac{499}{1000}$	$\frac{501}{1000}$
Resultados		$P(H_1 \mid R \wedge X)$	$P(H_2 \mid R \wedge X)$	$P(H_1 \mid R \wedge X)$	$P(H_2 \mid R \wedge X)$
1	ca	0,999	0,001	0,501	0,499
2	co	0,500	0,500	0,500	0,500
3	ca	0,999	0,001	0,501	0,499
4	co	0,500	0,500	0,500	0,500
5	ca	0,999	0,001	0,501	0,499

Assim como no caso anterior, mantém-se a equivalência qualitativa dos parâmetros de decisão independentemente da magnitude do viés, como pode-se observar linha a linha pela atribuição de probabilidades maiores às mesmas hipóteses.

Como a sequência responsável pelo pior caso é muito mais provável de ser obtida em caso de uma moeda justa, tem-se uma situação curiosa em que o agente, conforme modelado, apresenta uma alta sensibilidade a vieses nos casos em que eles existirem e, ao mesmo tempo, terem um desempenho médio igual ao obtido por meio de qualquer regra para o caso de tratar-se de uma moeda justa, ainda que essa opção sequer esteja representada como hipótese.

[4] Seguindo o princípio básico da racionalidade instrumental, um valor esperado positivo viabiliza uma aposta, um negativo a inviabiliza e um valor igual a zero a torna irrelevante.(Wallace, 2009)

Para não contrariar a resposta dada há pouco sobre a não-razoabilidade de se obter bons resultados usando hipóteses falsas, como parece implicar o que acaba de ser mostrado, vale lembrar que esse desempenho só é possível graças a todo um conjunto particular de características das situações específicas tratadas, fazendo com que a principal diferença em termos de conhecimento representado nas tabelas (a possibilidade de tratar-se de moeda justa) não permita nenhuma vantagem em termos de acerto de predição.

3.4 Considerações Finais

A pesquisa realizada tentou levar em conta tanto abordagens tradicionais quanto modernas, o que pode ser visto em 1.3.2. Adicionalmente, foi dada atenção especial aos conceitos de *Teoria de Probabilidades* e *Teoria de Conjuntos* como forma de aproveitar a forte conexão com o problema da indução, embora a aplicação a esse caso específico ainda seja relativamente rara, com exceção notável para a tese de Williams-Stove (Stove, 1986).

Para aplacar a curiosidade do leitor, pode-se adiantar que esse estudo se posiciona ao lado daqueles que defendem a racionalidade da indução.

Embora alguns dos argumentos desenvolvidos acabem por ser equivalentes ou reforçar a tese de Williams-Stove, o que é compreensível dado o compartilhamento do enfoque probabilístico, já há defesas específicas dele (Campbell and Franklin, 2004), assim como defesas do projeto de Carnap (Maher, 2010) e do conceito geral de probabilidade enquanto lógica (Franklin, 2001; Jaynes and Bretthorst, 2003).

Desse modo, a principal contribuição apresentada é a tentativa de estabelecer a solução para o **problema da indução** ainda no domínio da própria definição de racionalidade, considerando uma abordagem evolucionário-epistemológica (Bradie and Harms, 2012) com relação a regras de predição conforme mostrado em 3.3.1, estabelecendo a **aposta no evento em maior proporção** como o a regra racional geral da qual a decisão em situações de certeza é um caso especial, pelo menos no âmbito das restrições consideradas.

Aparentemente o fato de as demais abordagens assumirem a regra da maior proporção como premissa pode ser uma das explicações para a dificuldade do tratamento do problema, além da questão das definições de racionalidade e probabilidade mencionada em 3.3.

Ao fazer essa assunção, elas partem para os cálculos quantitativos das proporções e para a justificação desses cálculos dadas as informações disponíveis e os novos dados obtidos, processo que se encontra bastante desenvolvido em termos técnicos. Esse desenvolvimento nos deixa numa situação curiosa em que, uma vez justificada a racionalidade da premissa (apostar no evento em maior proporção), já se dispõe de argumentos indutivos partindo dela e aplicados a inúmeros campos de conhecimento, além de situações do senso comum.

Esses argumentos podem ser vistos como o processo de estabelecer inicialmente em que probabilidades se deve crer e em seguida como manipular essas probabilidades frente à aquisição de mais conhecimento, estendendo a aplicabilidade do critério de racionalidade considerado a casos onde há incerteza sobre as proporções, conforme visto em 3.3.

Uma vez que as observações aqui desenvolvidas partem de um conjunto particular de experimentos para argumentar em prol da indução probabilística como método de aplicabilidade geral, é possível que os problemas de circularidade descartados como referentes a tentativas específicas de justificação sejam levantados também contra a nossa abordagem, afinal de contas, ao apelarmos para modelos estamos de certa forma utilizando a experiência, ainda que apenas idealmente, para justificar a indução.

Ainda não temos uma resposta conclusiva a essa crítica, apesar de ela nos parecer menos grave pelo fato de a questão da racionalidade ter sido deslocada praticamente “para fora” do sistema de atribuição de probabilidades, recaindo sobre a questão normativa da crença (por que crer no mais provável?).

Além disso, uma outra defesa possível, ainda que se admita certa circularidade na forma da assunção de uma hipótese inicial a partir da qual seriam realizadas as previsões, é que o processo indutivo considerado possui mecanismos de controle que permitem o questionamento a respeito de seu próprio potencial preditivo. Vejamos, por exemplo, o caso da Mega-sena onde a hipótese favorecida pelos dados é a de que os resultados são imprevisíveis diante do conhecimento atual. Nesse caso, os próprios cálculos probabilísticos nos estão dizendo que a informação passada (mais o conhecimento disponível) não nos permite prever o futuro.

Deixando o enfoque geral, dois pontos específicos merecem ser considerados isoladamente: o primeiro deles sendo o exemplo ícone do problema da indução e o outro a famosa "solução explicativa" proposta pelo próprio Hume. Antes de passarmos diretamente a eles, fechemos com uma citação referente a uma das defesas da racionalidade da indução mencionadas anteriormente.

"Análise Bayesiana moderna é exatamente a única expressão quantitativa desse formato de raciocínio, o raciocínio indutivo que Hume e Popper supuseram impossível."[5] (Jaynes and Bretthorst, 2003, pg. 341, tradução nossa).

3.4.1 Sobre o nascer do sol

O exemplo do sol é particularmente interessante por representar um caso onde a nossa intuição demanda que apostemos com certeza em seu nascimento no dia de amanhã, o que parece tornar inadequada a aplicação do conceito de probabilidade (ainda que elevada) a esse evento, como o próprio Hume observou ao afirmar:

"Pareceria ridículo aquele que dissesse que é apenas provável que o sol nascerá amanhã ou que todos os homens são mortais. Apesar disso, é óbvio que não temos nenhuma certeza a respeito desses fatos, além daquela que a experiência nos permite."[6] (Hume, 1978, 1.3.11, tradução nossa).

Além dessa característica intuitiva, a importância que o sol representa para nosso planeta faz com que não haja a opção de "não apostar" em uma das possibilidades

[5] Modern Bayesian analysis is just the unique quantitative expression of this reasoning format, the inductive reasoning that Hume and Popper held to be impossible.

[6] One would appear ridiculous, who would say, that it is only probable the sun will rise to-morrow, or that all men must dye; though it is plain we have no further assurance of these facts, than what experience affords us.

quanto ao seu nascimento no dia de amanhã, já que o nosso comportamento padrão assume que ele nascerá, e se pensássemos diferente provavelmente agiríamos conforme alguma das inúmeras ilustrações artísticas das atitudes humanas frente a um abrupto fim do mundo.

Isso nos coloca em um caso similar ao das apostas necessárias dos exemplos considerados cuja equivalência foi representada, como visto em 3.2, pela adaptação dos questionamentos sobre as moedas para:

- Se tanto a hipótese da regularidade quanto a da não-regularidade do nascer do sol são compatíveis com os resultados obtidos, como escolher entre elas? Que tipo de impacto na plausibilidade dessas hipóteses pode haver, dado que o nascimento tem ocorrido até hoje?

- Especificamente sobre o próximo dia, que parâmetro considerar para escolher entre as opções disponíveis (sol nasce $\times$ sol não nasce), já que ambas são logicamente possíveis?

Uma vez que a analogia com o caso da moeda parece consistente, definidas as hipóteses iniciais, segue-se o mesmo procedimento probabilístico para atualizar as plausibilidades à medida que novos dados são levados em conta. Dessa forma, cada nascimento extra fornece mais evidência a qualquer hipótese de regularidade, o inverso acontecendo para o caso de um não-nascimento.

Quanto à segunda pergunta, uma vez que os nascimentos anteriores tenham tornado a hipótese de regularidade mais plausível, naturalmente a aposta no nascer do sol segue como escolha racional.

É importante observar mais uma vez que a escolha das hipóteses iniciais é diretamente proporcional à quantidade de informação que se deseja assumir como ponto de partida, tendo sido justamente a fonte de inúmeros questionamentos a respeito da famosa regra de sucessão de Laplace, que resolveu usar como exemplo o caso do nascer do sol, mas desconsiderando grande parte da informação naturalmente à nossa disposição, para fins ilustrativos.

Apesar de ter alertado explicitamente para esse fato, Laplace foi criticado mesmo por cientistas e filósofos simpáticos ao uso da teoria de probabilidades como ferramenta científica.[7]

3.4.2 Sobre o hábito

Retomando a questão do hábito sob um novo enfoque, coerente com as ideias expostas anteriormente, pode-se dizer que Hume foi bastante preciso quanto à conexão com as origens do raciocínio indutivo, principalmente se o entendermos enquanto atitude inconsciente embora deixando em aberto a possibilidade de justificação.

Isso nos colocaria de certa forma alinhados a Hume quanto à racionalidade no que concerne ao indivíduo, já que ao não ter consciência dos mecanismos de funcionamento do Hábito não podemos dizer sequer que a escolha por reprimi-lo ou permitir sua expressão poderia ser feita racionalmente.

Ao considerarmos a abordagem evolucionário-epistemológica juntamente com a definição de racionalidade conforme 3.3.1, além das formulações mais recentes dos fundamentos da teoria de probabilidades, parece ser possível não só diferenciar os hábitos com relação a critérios básicos de racionalidade instrumental como também identificar a convergência daqueles relacionados à efetivação de escolhas de modo a apostarem conforme as proporções dos eventos considerados. Essa convergência, mais visível à medida que as proporções se afastam de uma distribuição equitativa, já iria contra o prognóstico inicial de Hume, já que a situação considerada envolve incerteza, mas a performance de um tipo de regra é necessariamente melhor que a da outra se consideradas no âmbito geral da população de agentes envolvida (ou seja: um hábito é melhor que o outro).

Tomando as proporções como pressão evolutiva, por exemplo, pode-se pensar na racionalidade da escolha guiada pelas probabilidades como sendo externa ao agente primitivo e, portanto, ele não seria racional no sentido de entender as causas de seu hábito.

[7] Mais sobre a história e defesa da racionalidade da regra de sucessão em (Jaynes and Bretthorst, 2003).

Para um observador externo, entretanto, uma vez que o agente em questão tenha sido selecionado (a regra usada, seja qual for, está ajustada o suficiente para garantir sua sobrevivência), fica evidente que uma regra racional candidata a substituir a atual teria de gerar necessariamente apostas em proporções ainda mais próximas da real.

O próximo passo nesse processo de "racionalização do hábito" seria justamente aquele em que a visão do observador externo começa a ser internalizada pelo agente através da aquisição de mais informação sobre o ambiente, iniciando apenas com suas próprias experiências e depois passando a usar as experiências adquiridas por outros agentes em grupos cada vez maiores.

Por fim, pode-se resumir a proposta desse estudo para o status do hábito com relação ao processo de aquisição de conhecimento nos seguintes itens:

- A incerteza na escolha não impossibilita necessariamente a distinção entre hábitos baseada em critérios de racionalidade instrumental;
- Os hábitos, desde que as escolhas deles decorrentes tenham alguma influência em sua preservação, serão forçados a convergir em conformidade com as proporções entre os eventos em questão e, por isso, podem ser vistos como "controlados" pelas probabilidades;
- Um hábito, enquanto mecanismo de escolhas sob incerteza, será racional na medida em que se aproxima da aposta no evento em maior proporção. Um agente portador de tal hábito será racional na medida em que tem consciência dessa aproximação e, se confrontado com um hábito concorrente, faz a escolha baseado nesse parâmetro.

Pode-se dizer, então, que a racionalidade pensada em relação ao hábito nos permite adaptar a famosa frase de Hume para:

A razão é escrava da auto-replicação.

Referências Bibliográficas

Arnborg, S. and Sjödin, G. (2001). On the foundations of Bayesianism. In *AIP Conference Proceedings*, volume 568, page 61.

Audi, R. (1999). *The Cambridge Dictionary of Philosophy*. Cambridge University Press, New York, second edition.

BonJour, L. (2009). *Epistemology: Classic problems and contemporary responses*. Rowman & Littlefield Pub Incorporated.

Bradie, M. and Harms, W. (2012). Evolutionary Epistemology. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Spring 2012 edition.

Broad, C. D. (1952). *Ethics and The History of Philosophy*. Routledge.

Campbell, S. and Franklin, J. (2004). Randomness and the Justification of Induction. *Synthese*, 138(1):79–99.

Carnap, R. (1962). *Logical Foundations of Probability*. The University of Chicago Press.

Empiricus, S. (1933). *Outlines of Pyrrhonism*. Loeb Classical Library.

Floridi, L. (2011). Semantic Conceptions of Information. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Spring 2011 edition.

Forster, M. R. (2005). Notice: No Free Lunches for Anyone, Bayesians Included.

Franklin, J. (2001). Resurrecting logical probability. *Erkenntnis*, 55(2):277–305.

Futuyma, D. J. (2005). *Evolution*. Sinauer Associates, Inc.

Gensler, H. J. (2010). *Introduction to logic*. Routledge.

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis*, pages 121–123.

Goodman, H. N. (1983). *Fact, Fiction, and Forecast*. Harvard University Press.

Hamming, R. W. (1991). *The art of probability*. Addison Wesley Publishing Company.

Hawthorne, J. (2012). Inductive Logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2012 edition.

Hempel, C. G. (1981). Turns in the evolution of the problem of induction. *Synthese*, 46(3):389–404.

Hofweber, T. (2012). Logic and ontology. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2012 edition.

Howson, C. (2003). *Hume's Problem: Induction and the Justification of Belief*. Oxford University Press.

Hume, D. (1978). *A Treatise of Human Nature*. Oxford University Press.

Hume, D. (2007). An Enquiry Concerning Human Understanding. Obtido via Project Gutenberg em: http://www.gutenberg.org/ebooks/9662/.

Hájek, A. (2012). Interpretations of Probability. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2012 edition.

Jaynes, E. T. and Bretthorst, G. L. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

Jech, T. J. and Hrbacek, K. (1999). *Introduction to set theory*, volume 220. CRC Press.

Keynes, J. M. (1921). *A Treatise on Probability*. Macmillan & Co., Ltd.

Lehrer, J. (2009). *How we decide*. Houghton Mifflin Harcourt (HMH).

Maher, P. (2010). Explication of inductive probability. *Journal of philosophical logic*, 39(6):593–616.

Maxwell, J. C. (1990). *The Scientific Letters and Papers of James Clerk Maxwell: 1846-1862*, volume 1. Cambridge University Press.

Popper, K. R. (2002). *The Logic of Scientific Discovery*. Routledge.

Pritchard, D. and Turri, J. (2012). The value of knowledge. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Winter 2012 edition.

Ross, S. M. (2009). *Introduction to probability models*. Academic press.

Safire, W. (1993). ON LANGUAGE; Words Out in the Cold.

Savant, M. V. (1990). Ask Marilyn [column]. *Parade Magazine*, page 16.

Selvin, S. (1975). On the Monty Hall problem. *American Statistician*, 29(3):134.

Sewell, M. (2012). No Free Lunch Theorems. Endereço: http://www.no-free-lunch.org/.

Skinner, B. F. (1948). Superstition in the pigeon. *Journal of experimental psychology*, 38(2):168.

Steup, M. (2012). The analysis of knowledge. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2012 edition.

Stove, D. C. (1965). Hume, probability, and induction. *The Philosophical Review*, 74(2):160–177.

Stove, D. C. (1986). *The Rationality of Induction*. Oxford University Press.

Vickers, J. (2011). The Problem of Induction. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2011 edition.

Wallace, R. J. (2009). Practical Reason. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2009 edition.

Weintraub, R. (1995). What was Hume's Contribution to the Problem of Induction? *The Philosophical Quarterly*, 45(181):460–470.

Weirich, P. (2010). Causal Decision Theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2010 edition.

Wolpert, D. H. (1996). The lack of a priori distinctions between learning algorithms. *Neural computation*, 8(7):1341–1390.

Wolpert, D. H. (2012). WHAT DOES DINNER COST? Endereço: `http://www.no-free-lunch.org/coev.pdf`.

Appendices

Apêndice A

Programa coin-all.py

Programa simples construído conforme a assunção inicial das probabilidades às hipóteses possíveis, de forma a simular a realização do experimento e checar a adequação do cálculo do valor esperado.

```
# Resultado da regra de maior numero de ocorrencias para as 3 possibilidades de moedas
# Aposta inicial em coroa, atualizada para o resultado mais frequente (empate=coroa)

# Linhas para permitir uso de funcoes especificas
import time, random
from datetime import timedelta
from random import randint

start=time.clock() # inicio da contagem do tempo de execucao do programa (opcional)

# Inicializacao das variaveis (total = considerando todos os experimentos)
Headstotal=0 # Numero total de caras
Tailstotal=0 # Numero total de coroas
Hits=0 # Numero total de acertos de predicoes
Experiments=0 # Numero de experimentos a serem realizados

# Experimento (composto por 10 lancamentos)
while (Experiments<1000000000):
        Possibilities=randint(0,2)
        Heads=0
        Tails=0
        Trials=0
        while (Trials<10): # Realizacao dos 10 lancamentos
                if (Possibilities==0):
```

```
                toss=0
            elif (Possibilities==1):
                toss=1
            else:
                toss=randint(0,1)
            if (Heads>Tails):
                bet=0
            else:
                bet=1
            if (toss==0):
                Heads=Heads+1
                Headstotal=Headstotal+1
            else:
                Tails=Tails+1
                Tailstotal=Tailstotal+1
            if (bet==toss):
                Hits=Hits+1
            Trials=Trials+1
        Experiments=Experiments+1

stop=time.clock() # Fim da contagem do tempo de execucao

# Apresentacao dos resultados na tela
print 'Heads = ',Headstotal,'Tails = ',Tailstotal,'Hits = ',Hits,' | tempo: ',stop-start
```