



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"

7 a 9 de novembro de 2018
Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ



O Princípio da Máxima Entropia e o Problema da Razão Insuficiente

Marcelo Mattos Antunes
HCTE-UFRJ
e-mail seacelo@hotmail.com

Alexandre Lyra Oliveira
Observatório do Valongo-UFRJ
e-mail alexandr@astro.ufrj.br

Resumo: O Princípio da Máxima Entropia estabelecido em 1957 por E. Jaynes trata de generalizar para a Mecânica Estatística as leis de inferência estabelecidas a partir da Teoria das Comunicações de C. E. Shannon. Ao elaborar seu trabalho, Jaynes vinculou seu princípio ao denominado Princípio da Razão Insuficiente, o qual vinculou a Laplace. Nosso trabalho busca investigar epistemologicamente estes diferentes Princípios buscando, por exemplo, nos trabalhos de Bernoulli, Laplace, Keynes e outros, as raízes da Máxima Entropia.

Palavras-chave: Entropia, Probabilidade, Informação e Incerteza

1. A Teoria da Informação e o Princípio da Máxima Entropia

O conceito de entropia foi introduzido na Ciência no século XIX e alguns princípios que o envolvem têm suas origens nos estudos sobre a termodinâmica, mais efetivamente, com os trabalhos de Clausius, Maxwell, Boltzmann e Gibbs.

No século XX, esses estudos tornaram-se fundamentais para várias áreas do conhecimento científico e suas aplicações vão desde questões que envolvem a evolução humana, o clima, a Cosmologia até na reconstrução de imagens tomográficas.

No ano de 1948 o engenheiro norte americano Claude E. Shannon publicou, pela primeira vez, a sua Teoria Matemática da Informação. Nessa teoria, foi mostrado que a maioria dos canais de comunicação possui uma capacidade ou uma taxa de transmissão e que a informação só pode ser transmitida através do canal se,



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"

7 a 9 de novembro de 2018
Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ



e somente se, a quantidade de informação enviada pela fonte do canal não exceder a sua capacidade de transmissão. Por esse motivo, Shannon procurou otimizar os meios de comunicação daquela época, diagnosticando a capacidade de transmissão de cada canal e o nível de confiabilidade em cada informação enviada, desde o emissor até o receptor.

De acordo com Shannon, toda informação está associada à incerteza e a medida utilizada para quantificar essa incerteza é a mesma expressão matemática da entropia da termodinâmica, exceto pela presença da constante de Boltzmann.¹ Desse modo, o conceito de entropia, que anteriormente estava vinculado somente aos problemas da física, puderam ser aplicados a uma grande variedade de problemas, causando um grande impacto em diversas áreas do conhecimento científico daquela época, sobretudo para engenharia da comunicação.

Mas apesar de toda essa diversidade de problemas que puderam ser tratados pela teoria da informação, ainda não estava claro como essa teoria poderia ser aplicada à mecânica estatística e aos seus problemas subjacentes. Isso só foi possível com o físico americano Edwin Thompson Jaynes que, refletindo sobre as questões apontadas na teoria da informação, desenvolveu o Princípio da Máxima Entropia. Conforme Jaynes afirmou em seu artigo, *a teoria da informação forneceu um critério para atribuir distribuições de probabilidade com base num conhecimento parcial, que leva a um tipo de inferência estatística denominada estimativa de máxima entropia.*²

Atribuir probabilidades quando não há nenhuma informação disponível ou quando as informações são apenas parciais, foi um dos principais obstáculos para os trabalhos iniciais da termodinâmica. Por exemplo, em Boltzmann que, segundo Jaynes, abriu o caminho para o Princípio da Máxima Entropia, colocando as seguintes perguntas: I) De quantas maneiras diferentes um determinado número de moléculas pode ser distribuído? II) Entre todas as distribuições possíveis, qual é a

¹ É essencial observar, de acordo com Jaynes, que o mero fato de que a mesma expressão matemática $-\sum p_i \log p_i$ ocorre tanto na mecânica estatística quanto na teoria da informação não estabelece, por si só, qualquer ligação entre esses campos. Isso só pode ser feito sob novos pontos de vista de que *entropia da termodinâmica* e a *entropia da teoria da informação* aparecem com o mesmo conceito. Jaynes, E.T. "Information Theory and Statistical Mechanics". Department of Physics, Stanford University, Stanford, California (Received September 4, 1956; revised manuscript received March 4, 1957)

² O artigo de Jaynes sobre o Princípio da Máxima Entropia foi publicado pela primeira vez em 1957.



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"



7 a 9 de novembro de 2018
Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ

mais provável? A resposta de Boltzmann, que serviu como ponto de partida para Jaynes, foi que a distribuição "mais provável" é aquela que pode ser realizada pelo maior número de caminhos possíveis; isto é, aquela que pode ser maximizada sujeito a certas restrições³.

Seguindo essa linha de raciocínio, Jaynes concluiu que para fazer inferências com base em informações parciais, devemos usar a distribuição de probabilidade que tenha a máxima entropia a partir de qualquer informação conhecida. Na linguagem informal, trata-se de prever aquilo que não se vê.

Para refletir essa ideia, tomemos o seguinte exemplo: suponha que a única informação parcial disponível é que a temperatura média de uma cidade durante quatro dias consecutivos foi de 25°. Entretanto, não sabemos qual foi a temperatura em cada dia. De acordo com a média fornecida, temos várias distribuições possíveis, por exemplo, {27°, 26°, 22° e 25°}, {24°, 28°, 20° e 28°}, {25°, 25°, 25°, 25°} e {14°, 20°, 30° e 36°}. Essas são apenas algumas distribuições entre uma infinidade de configurações possíveis, que são compatíveis com a informação inicial disponível, ou seja, a média de 25°. Certamente a distribuição mais imparcial ou menos tendenciosa, que reflete o nosso desconhecimento inicial ou a nossa incerteza quanto à temperatura diária, é a distribuição mais uniforme possível, ou seja, {25°, 25°, 25°, 25°}. Nesse caso, como não há *nenhuma razão* ou informação inicial, *a priori*, que possibilite privilegiar algum resultado (a temperatura), deve-se atribuir o mesmo valor a todos os eventos possíveis. Essa situação, expressa, de certa forma, o "princípio da razão insuficiente" ⁴ que, como veremos no decorrer desse trabalho, foi motivo de reflexão para Jaynes em seu Princípio da Máxima Entropia.

Nesse sentido, o cerne do problema é obter um método de amostragem que não seja tendencioso e que esteja de acordo com qualquer informação disponível, que no caso do nosso exemplo foi a média.

³ "In Boltzmann's reasoning at this point, we have the beginning of the Principle of Maximum Entropy. He asked first: In how many ways could a given set of occupation numbers $\{N_k\}$ be realized? The answer is the multinomial coefficient [...]. Now any set $\{N_k\}$ of occupation numbers for which E, N agree with the given information, represents a possible distribution, compatible with all that is specified. Out of the millions of such possible distributions, which is most likely to be realized? Boltzmann's answer was that the "most probable" distribution is the one that can be realized in the greatest number of ways; Jaynes, E. T. "Where do we Stand on Maximum Entropy?", p.15. (nossa tradução).

⁴ Essa expressão foi colocada entre aspas para indicar que não se trata, formalmente, de um princípio.



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"

7 a 9 de novembro de 2018
Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ



Conforme Jaynes afirmou em seu Princípio da Máxima Entropia “se os resultados concordam ou não com a experiência, eles ainda representam as melhores estimativas que poderiam ser feitas com base nas informações disponíveis.” Na mesma obra ele acrescenta: “O processo de inferência de máxima entropia é aquele em que escolhemos a distribuição de probabilidade mais ampla possível sobre os estados microscópicos, compatível com os dados iniciais.”⁵

Concisamente podemos dizer de que o Princípio de Máxima Entropia possibilita tratar vários fenômenos que envolvem imprevisibilidade e incerteza, como ocorre frequentemente na análise de sistemas estatísticos, na mecânica quântica e na termodinâmica, onde um dos maiores obstáculos está na correspondência entre os fenômenos microscópicos e macroscópicos.

Certamente Jaynes retomou os problemas elementares da teoria da informação em seu princípio da máxima entropia, principalmente quando ele afirma “Que o grande avanço que a teoria da informação forneceu está na descoberta de que existe um critério único e inequívoco para "quantificar a incerteza" representada por uma distribuição discreta de probabilidade, o que está de acordo com nossas noções intuitivas de que uma ampla distribuição representa mais incerteza do que uma precisão acentuada, e satisfaz todas as outras condições que a tornam razoável” [2].

O critério que satisfaz as condições apontadas por Jaynes é o Princípio da Máxima Entropia, onde ele estabelece que para maximizar a entropia é preciso atribuir a distribuição de probabilidade que represente o máximo de incerteza, considerando certas condições previamente determinadas. Para isso, é necessário que a distribuição de probabilidade tenha o maior grau de semelhança entre suas probabilidades, ou que se aproxime mais de uma distribuição uniforme ($1/n$) distinguindo-se dela apenas pelas restrições ou por alguma informação inicial como a média da distribuição, por exemplo.

Do mesmo modo como ocorre na teoria da informação de Shannon, o conceito de entropia, em Jaynes, significa incerteza que, por sua vez, está associada a uma distribuição de probabilidade. Nesse caso, havendo mais eventos há mais escolhas possíveis e maior será a incerteza. Segundo Jaynes, “uma ampla

⁵ Jaynes, E.T. “Information Theory and Statistical Mechanics”.



distribuição de probabilidade representa mais incerteza do que uma precisão acentuada".⁶ Inserido nesse contexto, em particular no problema das probabilidades uniformes, encontra-se o "princípio da razão insuficiente", que discutiremos a seguir.

2. O "princípio da razão insuficiente" no Princípio da Máxima Entropia

O Princípio da Máxima Entropia está relacionado com o problema da especificação de probabilidades, ou seja: como atribuir probabilidade a um determinado evento quando não temos nenhuma informação disponível ou quando as informações são apenas parciais? Segundo Jaynes, esse problema é tão antigo quanto à própria teoria das probabilidades de Laplace, cujo único meio disponível para atribuir probabilidades, nessas condições, foi o controverso "Princípio da Razão Insuficiente", que Jaynes interpreta e da seguinte forma: "Dois eventos devem receber probabilidades iguais se não há razão para pensar ao contrário." ⁷ Entretanto, esse argumento só se justifica quando há uma clara evidência ou, de acordo com Jaynes, uma simetria indicando que todos os eventos são igualmente possíveis, esta hipótese pode parecer tão arbitrária quanto qualquer outra que poderia ser feita. Por esse motivo, essa maneira de abordar esses tipos de problemas foi abandonada, devido à ausência de qualquer princípio construtivo que nos desse uma razão para preferir uma probabilidade distribuição em relação à outra em casos em que ambas têm as mesmas informações disponíveis.

Segundo Jaynes, para uma discussão mais aprofundada deste problema, deve-se reconhecer o fato de que a teoria da probabilidade foi desenvolvida em duas direções muito diferentes, uma seguindo a escola de pensamento "objetivo" e outra a linha do pensamento "subjetivo". Nesse contexto, Jaynes coloca duas perguntas fundamentais para teoria das probabilidades: o teste de uma boa distribuição de probabilidade objetiva $p(x)$ é aquela que representa corretamente todas as variações x ? O teste de uma boa distribuição de probabilidade subjetiva é

⁶ Jaynes, E. T. "Prior Probability", IEEE Transactions On Systems Science and Cybernetics, vol.4 sec.4 N.3 (1968).

⁷ Utilizamos aspas acompanhando o texto original de Jaynes e também para indicar que essa denominação "princípio da razão insuficiente" não foi utilizada nem por Laplace e sim por seus sucessores, como Jaynes por exemplo.



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"

7 a 9 de novembro de 2018
Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ



aquela que representa corretamente nosso estado de conhecimento quanto ao valor de x ?

Jaynes, respondendo a primeira pergunta, considera que a probabilidade objetiva de um evento corresponde a uma medida dada por frequências relativas de um experimento aleatório. Desse modo, no cálculo de uma distribuição de probabilidade objetiva, acredita-se que as previsões são, em princípio, verificáveis detalhadamente e não dependem de nenhuma opinião pessoal.

Por outro lado, a escola do pensamento "subjetivo" considera as probabilidades como expressões da ignorância humana⁸ ou da expectativa de que um evento poderá ocorrer ou não, com base em alguma informação que esteja disponível. Desse modo, no cálculo de uma distribuição de probabilidade subjetiva, acredita-se que as previsões são independentes de qualquer processo experimental e dependem apenas de alguma informação particular que esteja disponível.

Em alguns tipos de sistemas a quantidade de informação disponível é tão pequena que não se pode atribuir probabilidades por meio de experiências e, nesse caso, não se pode fazer previsões confiáveis. Nesse sentido, Jaynes considera a visão subjetiva é mais abrangente do que a visão objetiva, pois possibilita atribuir probabilidades a partir de qualquer informação que esteja disponível.

3. Considerações finais

O "Princípio da razão Insuficiente" foi alvo de várias críticas, especialmente aquelas que questionavam a sua validade como um critério para atribuir probabilidades, considerando que não havia nenhuma fundamentação teórica ou filosófica que pudesse justificar a noção de uma razão insuficiente como um princípio. Entre os críticos desse "princípio" encontram-se alguns nomes de peso para a teoria das probabilidades, como o de G. Boole, J. Venn, Fisher e K. Pearson.⁹ Por outro lado, esse princípio foi retomado nas obras de Keynes, Jeffreys e Jaynes,

⁸ Keynes, J. M. "Treatise on Probability" - MacMillan Company, London, 1921. Citado por Jaynes em "Information Theory and Statistical Mechanics I" [2].

⁹ As críticas relacionadas ao nome "princípio da razão insuficiente" são variadas, mas investigá-las com mais profundidade não está nos objetivos do presente trabalho. Maiores detalhes sobre essas críticas podem ser encontrados no artigo de J. Uffink, "Can the maximum entropy principle be explained as a consistency requirement?", 1997.



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"

7 a 9 de novembro de 2018
Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ



que considerou o Princípio da Máxima Entropia como uma generalização do "princípio da razão insuficiente".

Conforme mencionamos anteriormente, Jaynes segue a linha subjetiva de probabilidades. Nesse sentido, o problema de uma razão insuficiente envolve, necessariamente, um caráter subjetivo de julgamento quanto à suficiência ou insuficiência de nossa razão. Além disso, o que é razoável para um determinado ponto de vista pode não ser para outro. Desse modo, torna-se questionável falar de "ausência de razão suficiente" sem envolvimento de juízos subjetivos, pois essa condição reflete um estado de conhecimento que pode ser vago ou incompleto. Por exemplo, se uma pessoa vai assistir uma corrida de carros esportivos e desconhece a capacidade de cada piloto, a potência de cada veículo e outros fatores técnicos, terá muita incerteza para estimar um possível vencedor. Nesse caso, a única informação disponível é que 25 carros irão competir, logo é razoável dizer, nessas condições, que todos os carros têm a mesma probabilidade de vencer, pois não há ou nenhuma razão ou nenhuma diferença para que um deles ganhe em vez de outro qualquer. Porém, se alguém acompanha o dia a dia de cada piloto e conhece a potência de cada carro, a situação muda completamente, ou seja: é possível, de acordo com essas condições disponíveis, distinguir a probabilidade que cada competidor tem para vencer a corrida. Esse exemplo mostra o caráter subjetivo que envolve a idéia de uma razão insuficiente.

De acordo com Laplace, "a probabilidade está relacionada em parte com nossa ignorância em parte ao nosso conhecimento. Sabemos que entre vários eventos, apenas um deve ocorrer, mas nada nos leva a crer que um deles ocorrerá preferencialmente aos outros. Nesse estado de indecisão, é impossível pronunciarmos com certeza sobre sua ocorrência." ¹⁰

Reconhecemos que vários historiadores das ciências atribuíram a Laplace a autoria do "princípio da razão insuficiente". Entretanto, nossos estudos indicam que a noção de uma razão insuficiente não corresponde a nenhum dos dez princípios

¹⁰ Laplace, P.S. *Essai Philosophique sur les Probabilités*, (published in 1814) Paris, Gauthier-Villars et editors, Troisième Édition, 1820.



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"

7 a 9 de novembro de 2018
Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ



estabelecidos por ele em suas obras *Ensaio Filosófico sobre as Probabilidades* [5] e *Teoria Analítica das Probabilidades* [6].

. A partir dessas investigações sugerimos que a noção de uma razão insuficiente encontra-se apenas subentendida na obra de Laplace como um argumento para atribuir probabilidades iguais entre vários eventos possíveis. Nossos estudos também indicam que outros cientistas que o precederam, já tinham utilizado esse mesmo argumento em suas obras de probabilidades, como Pascal, Huygens, Leibniz e Bernoulli.

Com esse estudo, procuramos identificar as origens históricas do problema da razão insuficiente e mostrar seus possíveis vínculos com o Princípio da Máxima Entropia de Jaynes.

REFERÊNCIAS

- [1] COUTURAT, L. "The Logic of Leibniz", from unpublished documents, Paris, Félix Alcan, 1901, reprinted Hildesheim, 1969.
- [2] JAYNES, E. T. Jaynes. "Information Theory and Statistical Mechanics I", Phys. Rev., Vol. 106, p. 620 and Vol. 108, p.171, 195
- [3] JAYNES, E. T. Jaynes. "Where do we stand on maximum Entropy", The Maximum Entropy Formalism, A Conference Held at the Massachusetts Institute of Technology on May 24, 1978. Edited by Raphael Do Levine and Myron Tribus.
- [4] KEYNES, J. M. Keynes, " A Treatise on Probability", 1921, Macmillan and Co. pp.41-64;
- [5] LAPLACE, P.S. "Philosophical Essay on Probabilities", (published in 1814) Paris, Gauthier-Villars et cie, Editors, 1921.
- [6] LAPLACE, P.S. "Analytical Theory of Probabilities". (published in 1812) ,Paris, Gauthier-Villars and Co., Publishers, 1921.
- [7] LEIBNIZ, G. W. , "L'estime des apparences". 21 manuscrits de Leibniz sur les probabilités, la théorie des jeux, l'espérance de vie. Texte établi, introduced par M. Parmentier, Vrin, Paris, 1998.
- [8] LEIBNIZ, G. W., "New Essays on Human Understanding". V. Civita, S. P., 1974.



Scientiarum Historia XI

Filosofia, Ciências e Artes: Conexões Interdisciplinares
"Sacudindo a poeira"



7 a 9 de novembro de 2018

Rio de Janeiro - RJ
HCTE/NCE/CCMN - UFRJ

[9] MARTYUSHEV, L. M., 2006, Maximum entropy production principle in physics chemistry and biology, 426, 1-45.

[10] RAYMOND, P. "De la Combinatoire aux Probabilités". Éditeur François Maspero, 1, place Paul-Painlevé, Paris-5°, 1975.

[11] SHANNON, C. E. Bell System Tech. J. 27, 379, 623 (1948); these papers are reprinted in C. E. Shannon and W. Weaver, "The Mathematical Theory of Communication" (University of Illinois Press, Urbana, 1949).

[12] UFFINK, Jos. "Can the maximum entropy principle be explained as a Consistency requirement?" Department of History and Foundations of Mathematics and Science (University of Utrecht, 1997).