UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Programa de Pós Graduação em História das Ciências e das Técnicas e Epistemologia

MARCELO MATTOS ANTUNES

ALGUMAS QUESTÕES EPISTEMOLÓGICAS DO PRINCÍPIO DA MÁXIMA ENTROPIA

RIO DE JANEIRO-RJ

2019

MARCELO MATTOS ANTUNES

ALGUMAS QUESTÕES EPISTEMOLÓGICAS DO PRINCÍPIO DA MÁXIMA ENTROPIA

Tese de Doutorado apresentada ao Programas de Pós-Graduação em História das Ciências e das Técnicas e Epistemologia da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Doutor em História das Ciências e das Técnicas e Epistemologia.

Orientador: Prof. Dr. Alexandre Lyra de Oliveira

RIO DE JANEIRO

2019

CIP - Catalogação na Publicação

M444a

Mattos Antunes, Marcelo
ALGUMAS QUESTÕES EPISTEMOLÓGICAS DO PRINCÍPIO DA MÁXIMA ENTROPIA / Marcelo Mattos Antunes. -- Rio de Janeiro, 2019.
81 f.

Orientador: Alexandre Lyra Oliveira.
Tese (doutorado) - Universidade Federal do Rio de Janeiro, Decania do Centro de Ciências Matemáticas e da Natureza, Programa de Pós-Graduação em História das Ciências e das Técnicas e Epistemologia, 2019.

1. Entropia. 2. Probabilidade. 3. Indiferença. 4. Simetria. 5. Incerteza. I. Lyra Oliveira, Alexandre, orient. II. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

MARCELO MATTOS ANTUNES

ALGUMAS QUESTÕES EPISTEMOLÓGICAS DO PRINCÍPIO DA MÁXIMA ENTROPIA

Tese de Doutorado apresentada ao Programas de Pós-Graduação em História das Ciências e das Técnicas e Epistemologia da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Doutor em História das Ciências e das Técnicas e Epistemologia

Aprovado em: 17 de dezembro de 2019

Alexandre Lyra de Oliveira, Dr.Sc., Orientador
Observatório do Valongo CCMN/UFRJ e HCTE/UFRJ

Zulena dos Santos Silva, Dra. Sc.,
Colégio Pedro II (CPII)-Departamento de Filosofia

Leandro L. S. Guedes, Dr. Sc.,
Fundação Planetario da Cidade do Rio de Janeiro

José Antonio dos Santos Borges, Dr. Sc.,
Instituto Tércio Pacitti / UFRJ e HCTE

Carlos Benevenuto Guisard Koehler, Dr. Sc.,
Instituto de Química / UFRJ e HCTE / UFRJ

Rundsthen V. Nader, Dr.Sc.
Observatório do Valongo CCMN/UFRJ e HCTE/UFRJ

DEDICATÓRIA

Aos meus filhos Patryck Berçot Antunes e Vanessa Parada Antunes pelos constantes incentivos para realizar esse trabalho.

A minha esposa, Rosane Ouriques Berçot Antunes, pela compreensão e companheirismo nos momentos de maior dificuldade.

Ao meu pai, Jorge Antunes, por me orientar como cidadão e profissional.

AGRADECIMENTOS

Ao meu orientador, Dr. **Alexandre Lyra**, pelo incentivo e pelas várias sugestões que contribuíram na elaboração desta tese.

Ao professor **Carlos Koelher** que, com suas aulas estimulantes de História das Ciências, me motivou desde o início para concluir o curso de doutorado.

Agradeço a Dra. **Zulena Silva** pelas preciosas discussões filosóficas.

Ao meu amigo astrônomo Dr. **Leandro Guedes** por seu incentivo na reta final deste trabalho.

Ao Secretário **Robson** por sua peculiar atenção e paciência com o corpo discente do HCTE.

“*Para melhor julgar sobre as pequenas percepções que somos incapazes de distinguir em meio à multidão delas, costumo utilizar o exemplo do bramido do mar, que nos impressiona quando estamos na praia. Para ouvir este ruído como se costuma fazer, é necessário que ouçamos as partes que compõe este todo, isto é, os ruídos de cada onda, embora cada um desses pequenos ruídos só se faça ouvir no conjunto confuso de todos os outros conjugados, isto é, no próprio bramir, que não se ouviria se esta onda que o produz estivesse sozinha”*

Leibniz

RESUMO

ANTUNES, Marcelo Mattos. **Algumas Questões Epistemológicas do Princípio da Máxima Entropia**. Rio de Janeiro, 2019. Tese (Doutorado em História das Ciências) – Programa de Pós-Graduação em História das Ciências, das Técnicas e Epistemologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

O Princípio da Máxima Entropia foi estabelecido por Jaynes em 1957 como um critério para obter distribuição de probabilidades nos casos em que há pouca ou nenhuma informação disponível. Segundo Jaynes, esse critério é uma generalização do princípio da razão insuficiente (ou princípio da indiferença), pelo qual se afirma que dois ou mais eventos devem ter probabilidades iguais se não houver nenhuma razão para se pensar ao contrário. Geralmente, o argumento desse princípio está vinculado ao nome de Laplace ou, em alguns casos, ao nome de Jacobi Bernoulli. Entretanto, um argumento semelhante a esse já se encontrava antes nos manuscritos de probabilidades de Leibniz. Portanto, pretende-se mostrar nesta tese que Leibniz foi um dos precursores dos conceitos que contribuíram para fundamentação do Princípio da Máxima Entropia. Para alcançar esse objetivo, procurou-se inicialmente revelar como os princípios e os conceitos, referentes ao problema da atribuição de probabilidades iniciais, foram estabelecidos originalmente nos trabalhos de probabilidades de Bernoulli e de Laplace, a fim de compará-los com os princípios e conceitos estabelecidos por Leibniz em seus manuscritos de probabilidades e também em algumas de suas principais obras filosóficas. Entre esses conceitos encontra-se o da indiferença de equilíbrio, que Leibniz enfaticamente rejeita, pois leva a ideia de uma escolha feita ao acaso ou sem nenhuma razão. Por outro lado, há uma indiferença, que ele admite, pela qual pode-se escolher livremente, entre vários casos possíveis, desde que essa escolha seja conduzida por uma razão determinante. O conceito de indiferença em Leibniz possibilitou, entre outras questões, analisar o problema da simetria que, de acordo com Jaynes, é uma condição necessária para atribuir probabilidades iniciais pelo princípio da indiferença.

Palavras-chave: Entropia. Probabilidade. Escolha. Indiferença. Simetria.

ABSTRACT

ANTUNES, Marcelo Mattos. Algumas Questões Epistemológicas do Princípio da Máxima Entropia. Rio de Janeiro, 2019. Tese (Doutorado em História das Ciências) – Programa de Pós-Graduação em História das Ciências, das Técnicas e Epistemologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

The Maximum Entropy Principle was established by Jaynes in 1957 as a criterion for obtaining probability distributions in cases where little or no information is available. According to Jaynes, this criterion is a generalization of the principle of insufficient reason (or the principle of indifference), which states that two or more events must have equal probabilities if there is no reason to think otherwise. Generally, the argument of this principle is linked to the name of Laplace or, in some cases, to the name of Jacobi Bernoulli. However, an argument similar to this was already found in Leibniz's probability manuscripts. Therefore, it is intended to show in this thesis that Leibniz was one of the forerunners of the concepts that contributed to the foundation of the Maximum Entropy Principle. To achieve this objective, we initially tried to reveal how the principles and concepts, referring to the problem of attributing initial probabilities, were originally established in the works of probabilities by Bernoulli and Laplace, in order to compare them with the principles and concepts established by Leibniz in his probability manuscripts and also in some of his major philosophical works. Among these concepts is the indifference of equilibrium, which Leibniz emphatically rejects, as it leads to the idea of a choice made at random and without any reason. On the other hand, there is an indifference, which he admits, whereby one can freely choose among several possible cases, provided that this choice is driven by a determining reason. The concept of indifference in Leibniz made it possible, among other issues, to analyze the problem of symmetry which, according to Jaynes, is a necessary condition for assigning initial probabilities by the indifference principle.

Keywords: Entropy. Probability. Choice. Indifference. Symmetry.

LISTA DE ILUSTRAÇÕES

SUMÁRIO

1 INTRODUÇÃO

Desde sua criação o conceito de entropia de Clausius tornou-se parte da termodinâmica e, posteriormente, da Mecânica Estatística e dos estudos de sistemas em equilíbrio e fora de equilíbrio. Sabemos que a bibliografia sobre esses temas é vastíssima e, como exemplo, podemos destacar: a *Produção de Entropia Máxima* de Dewar (2005); a *Produção de Entropia Mínima*[1] de Prigogine (1967, 1978) e o *Princípio da Máxima Entropia* de Jaynes (1957). Derivações dos dois primeiros princípios e do Princípio da Máxima Entropia (doravante MaxEnt), podem ser encontradas na literatura, por exemplo, com Martyushev (2006, p.46, et al). Dewar e Maritan (2014, p.49), enfatizam que o MaxEnt de Jaynes (1957), em sua formulação da mecânica estatística, fornece uma base teórica para o princípio da produção máxima de entropia. Contudo, há autores que restringem o MaxEnt quanto as suas aplicações. Esse é o caso de Shimony (1985) que divide as investigações sobre MaxEnt em dois grupos. Segundo ele, há autores que defendem e aplicam entusiasticamente o MaxEnt e, por outro lado, há autores que criticam e são céticos em relação a esse princípio. Uma parte dessas críticas foram logo respondidas pelo próprio Jaynes (1989, p.149), através de uma descrição minuciosa do MaxEnt desde seus fundamentos até as suas aplicações. Devemos mencionar também que foram escritos trabalhos exatamente em defesa do princípio de Jaynes, como é o caso de Tikochinsky (et al, 1984, p.357), que afirma: "O único algoritmo consistente é aquele que leva a distribuição de entropia máxima". Existem outros trabalhos que apontam alguns pontos a favor e também contra o MaxEnt além de fornecerem diversas referências sobre esse assunto, como é o caso de Pontzen e Governato (2013, p.121-133) que usam esse princípio para investigar a distribuição de matéria escura.

O MaxEnt é aplicado em várias áreas da física e também da astrofísica. Como exemplo, podemos destacar a análise espectral de Ables (1974, p.383), cujo "método produz representações espectrais superiores quando comparado com métodos tradicionais" e também fornece uma poderosa técnica de reconstrução de imagem como se vê em Skilling e Bryan (1984, p.111-124). No mesmo trabalho também encontramos outras aplicações do MaxEnt na Astronomia. No artigo de Gull e Daniell (1978, p.686-690) o MaxEnt é aplicado em astronomia de raios-X e

[1] O trabalho de Jaynes (1980) *The Minimum Entropy Production Principle* trata especificamente deste princípio.

também radioastronomia. Além disso, esse método também é aplicado na reconstrução e restauração de imagens tomográficas de raios-X, conforme constatamos no artigo de Mohammad e Demoment (1988, p.195). No caso da Astrofísica e da Cosmologia, também temos os trabalhos de Zunckel e Trotta (2007, p.865) que usam o MaxEnt na equação de estado da energia escura. Na Gravitação, com a confirmação em 2016 da existência das ondas gravitacionais previstas por A. Einstein, o estudo dos buracos negros assumiu ainda maior importância. Lembremos que estas primeiras detecções foram precisamente de colisões de buracos negros como nos mostrou Abott et al. (2016).

Recentemente publicamos (ANDREI et al. 2019, p.183-190) uma aplicação do MaxEnt na Astrofísica onde obtemos a distribuição da função de luminosidade nos quasars em diferentes redshifts, cuja previsão feita pelo MaxEnt é excelente sendo comparada com os dados observacionais. A fórmula desta distribuição, além de ser bastante simples, é totalmente nova na literatura.

Conforme destacou Jaynes (1989, p. 240-241), as investigações sobre os fundamentos do MaxEnt podem contribuir para que futuramente possam se estabelecer novos princípios[2] para diferentes tipos de informações testáveis e também para informações não-testáveis. No anexo A destacamos um exemplo com esses dois tipos de informações.

Além disso, várias questões epistemológicas e suas origens históricas puderam ser discutidas a partir do MaxEnt. Algumas dessas questões foram investigadas em nosso trabalho “The Genesis of the "Principle of Insufficient Reason" in leibnizian Thought and its Implications in the Principle of Maximum Entropy” (ANTUNES et al. 2017), que foi apresentado no 25th International Congress of History of Science and Technology. Posteriormente, em 2018, apresentamos no XI Scientiarum do HCTE/UFRJ “*O Princípio da Máxima Entropia e o Princípio da Razão Insuficiente* (ANTUNES; LYRA, 2018). Portanto, é inegável a importância do MaxEnt para Física e também para outras áreas do conhecimento científico.

Segundo Jaynes, (1989, p.9, grifo do autor) “o princípio da máxima entropia pode ser considerado como uma extensão do princípio da razão insuficiente”[3]. Esta generalização fica bem explicitada em sua formulação matemática pois, na medida

[2] Mais especificamente Jaynes (1989, p. 240-241) afirma que: “Talvez no futuro outros princípios podem ser descobertos, pelos quais as informações prévias não testáveis possam ser usadas em uma teoria matemática de inferência.

[3] Segundo Jos Uffink (1995, p.226), a origem desse nome é desconhecida. Geralmente essa expressão é colocada entre aspas para indicar que não se trata, formalmente, de um princípio e que não se sabe quem é o autor dessa denominação.

em que cada vínculo é adicionado à expressão da distribuição de probabilidade do MaxEnt, o valor da entropia diminui. Em outros termos, na medida em que obtemos mais informações sobre um sistema a distribuição de probabilidade torna-se cada vez mais precisa e, consequentemente, a incerteza[4] e a entropia diminuem proporcionalmente. Por outro lado, a entropia é máxima para o caso de um único vínculo, que é o da normalização das probabilidades, isto é, $\sum p_i = 1$. Este vínculo, isoladamente, leva à distribuição de probabilidade uniforme ($p_i = 1/n$) que, como será visto adiante, prova a consistência desse princípio como uma generalização do princípio da razão insuficiente.

Uma das principais questões epistemológicas abordadas no MaxEnt (1957) é o problema da atribuição de probabilidades que está vinculado ao princípio da razão Insuficiente e ao nome de Laplace:

> O '*Princípio Razão Insuficiente'* de Laplace foi uma tentativa de fornecer um *critério de escolha,* no qual se dizia que dois eventos devem ter *probabilidades iguais* se não houver nenhuma razão para se pensar ao contrário. Contudo, exceto nos casos onde há um elemento evidente de *simetria* que indica claramente que os eventos são "igualmente possíveis", esta hipótese pode parecer tão arbitrária quanto qualquer outra que poderia ser feita." (JAYNES, 1957, p.622, grifo do autor, tradução nossa) [5]

Notamos nessa passagem que há várias questões que podem ser discutidas a partir do ponto de vista do pensamento leibniziano, a saber, o problema da simetria, da escolha, da simetria e da equiprobabilidade. Como veremos mais adiante, essas questões encontram-se envolvidas no *conceito de indiferença* de Leibniz (1646-1716).

Posteriormente, nos fundamentos do MaxEnt e também nos antigos princípios das probabilidades, Jaynes vinculou o problema da atribuição de probabilidades iniciais a outros matemáticos:

> A base subjacente a essas atribuições iniciais foi declarada, explicitamente, como um princípio formal em Ars Conjectandi de Jacob Bernoulli (1713). Infelizmente, esse princípio recebeu o nome curioso de Razão Insuficiente que teve, desde então, uma certa rejeição que impede que muitos vejam o lado positivo dessa ideia. Keynes (1921) ajudou um pouco ao renomeá-lo como Princípio da Indiferença; mas até então o dano já havia sido feito. Se Bernoulli

[4] De acordo com Jaynes (1957, p.629), "a expressão da entropia é, literalmente, uma medida da quantidade de incerteza representada por uma distribuição de probabilidade."
[5] Laplace's "Principle of Insufficient Reason" was an attempt to supply a criterion of choice, in which one said that two events are to be assigned equal probabilities if there is no reason to think otherwise. However, except in cases where there is an evident element of symmetry that clearly renders the events "equally possible," this assumption may appear just as arbitrary as any other that might be made.

> tivesse chamado seu princípio, mais apropriadamente, de Desideratum de Consistency, ninguém teria tentado depreciá-lo. (JAYNES, 1989, p.212, tradução nossa) [6]

Essa passagem nos mostra que o problema da atribuição de probabilidades iniciais que, tecnicamente são chamadas de probabilidades anteriores[7], recebeu três denominações diferentes. O primeiro deles é o princípio da razão insuficiente que suscitou várias polêmicas de natureza técnica ou filosófica. Posteriormente Keynes (1921, p.44), o denominou como princípio da indiferença[8]. Já o próprio Jaynes, como vimos acima, sugere que o nome mais adequado ao problema da atribuição de probabilidades inicias é "Desideratum of Consistency" [9]. Entretanto, compreendemos que o mais importante nessa afirmação não está na discussão de um nome para um princípio, mas sim sobre sua essência que é, segundo Jaynes (1978, p. 213), reconhecer que a atribuição de probabilidade é um meio de descrever um certo estado de conhecimento e se a informação disponível não é suficiente para considerar uma proposição mais ou menos provável do que outra qualquer a maneira mais honesta de descrevermos esse estado de conhecimento é atribuir-lhes probabilidades iguais.

Até aqui vimos que Jaynes relacionou o problema da atribuição de probabilidades iniciais aos nomes de Laplace e Bernoulli. Entretanto, não há nenhuma referência ao nome de Leibniz e, conforme veremos mais adiante, algumas questões que fizeram parte da teoria das probabilidades de Bernoulli e de Laplace já estavam presentes nos manuscritos de probabilidades de Leibniz. Entre essas questões encontra-se a ideia inicial de equiprobabilidade ou de uma distribuição de probabilidades uniformes[10] que, naquela ocasião, era justificada pelo princípio da indiferença. Como já mencionamos anteriormente, argumenta-se

---

[6] The basis underlying such initial assignments was stated as an explicit formal principle in the Ars Conjectandi of Jacob) Bernoulli (1713). Unfortunately, it was given the curious name: Principle of Insufficient Reason which has had, ever since, a psychologically repellant quality that prevents many from seeing the positive merit of the idea itself. Keynes (1921) helped somewhat by renaming it the Principle of Indifference; but by then the damage had been done. Had Bernoulli called his principle, more appropriately, the Desideratum of Consistency, nobody would have ventured to deprecate it.

[7] Consistency requires it to recognize the relevance of prior information, and so in almost every problem it is faced at the onset with the problem of assigning initial probabilities, whether they are called technically prior probabilities or sampling probabilities. (JAYNES, 2003, p.343, tradução nossa).

[8] Heidelberger (2001, página 179) declara que Keynes, inspirado por von Kries, usou essa expressão como uma nova terminologia para o PRI de Laplace.

[9] Para maiores detalhes sobre essa denominação consultar *The Logic of Science*, Jaynes p. 17, 2003. Papers on Probability, Statistics and Statistical Physics, p. 210, 1983 e Jos Uffink, p.232, 1996 Can the Maximum Entropy Principle be Explained as a Consistency Requirement?

[10] Sobre esse assunto consultar o manuscrito intitulado *Sur le Calcul des Partis,* Leibniz (1995, p. 129)

por esse princípio que dois ou mais eventos devem ter probabilidades iguais se não houver nenhuma razão para se pensar ao contrário.[11]

Sabemos também que alguns dos antecessores de Leibniz já tinham tratado de questões semelhantes a essas em seus trabalhos de probabilidades, como é o caso de Pascal (1623-1662) que já pensava num critério para dividir, equitativamente, o valor de uma aposta entre dois jogadores.

> Para entender as regras das divisões, deve-se considerar primeiramente que o dinheiro apostado já não pertence mais aos jogadores. De acordo com as condições iniciais, eles poderão receber em troca tudo o que o acaso lhes der. [...] se um jogador ganhar, uma certa soma lhe pertencerá, e se perder, pertencerá ao outro. Como se trata de um jogo de puro acaso, *não há razão* para afirmar que um jogador vencerá e o outro não. (Pascal, p.460, 1952, grifo nosso)

Também pode-se encontrar exemplos semelhantes a esses em Huygens[12] (1629-1695). Entretanto, foi através na filosofia de Leibniz onde encontramos os elementos necessários para discutir os problemas correlatos aos fundamentos do MaxEnt como é o caso do princípio da indiferença. Mas antes de iniciarmos nossa discussão sobre esse princípio, reservamos parte dessa introdução para apresentar alguns fatos históricos que consideramos necessários para situar Leibniz no contexto histórico do desenvolvimento do cálculo das probabilidades. Apesar de alguns desses fatos históricos serem bem conhecidos, nos parecem relevantes contextualizá-los para identificar os possíveis enlaces do seu pensamento probabilístico com seus antecessores, por exemplo, com Huygens que teve grande importância na formação de Leibniz como matemático.

Conforme afirmou Jaynes (1978, p.210) a teoria da probabilidade começou com os *Ludo aleae* de Gerolamo Cardano em meados do século dezesseis. Além disso, o cálculo das probabilidades também recebeu um grande impulso com as correspondências trocadas entre os matemáticos franceses Fermat (1607-1665) e Pascal (1623-1662) e, respectivamente, com suas obras *Varia opera mathematica*, publicado em 1679 e o *Traité Du Triangle Arithmétique* de 1665.[13]

---

[11] Nesse caso estamos descrevendo a ideia básica desse princípio, visto que seu enunciado pode variar de acordo com a abordagem de cada autor.

[12] Huygens, C.: *Oeuvres complètes*, published by the Société Hollandaise des Sciences, 22 vols., The Hague, 1888-1950. Cf. Obra citada em "The Origins of the Infinitesimal Calculus". Baron, M. Oxford, Pergamon, 1969.

[13] Para maiores detalhes sobre as correspondências trocadas entre Pascal e Fermat consultar Hald (1990) *History of Probability and Statistics and Their Applications before 1750* e Edwards (1987) *Pascal's arithmetical triangle.*

Christian Huygens escreveu um tratado sobre jogos de azar com o título *De Ratiociniis in Ludo Aleae.*[14] Segundo Raymond (1975, p.113, tradução nossa), “a grande contribuição de Huygens está na composição de um tratado pedagógico, na explicitação da noção de chance e na extensão metodológica das soluções de Pascal.”[15]

Na citação que se segue, Leibniz nos mostra que já estava ciente de algumas dessas obras.

> Os matemáticos do nosso tempo começaram estimar os acasos durante os jogos. O Chevalier de Méré do qual, publicaram os *Ágréments* e outras obras, um homem de espírito penetrante, que era jogador e filósofo, deu uma oportunidade a isso, fazendo perguntas sobre o problema das partes, para descobrir quanto valeria o jogo se fosse interrompido em um determinado momento. Dessa maneira, ele conduziu o Sr. Pascal, seu amigo, para examinar um pouco essas coisas. A questão tornou-se conhecida e deu ao Sr. Huygens a oportunidade de fazer seu tratado sobre Alea. Outros homens instruídos também se interessaram por esse assunto. Assim, estabeleceram-se alguns princípios, que também foram utilizados pelo Sr. Witt em um pequeno discurso impresso em holandês, sobre as rendas vitalícias. (LEIBNIZ, 1974, p. 335)

Outra obra de grande importância para o desenvolvimento do cálculo das probabilidades foi *Ars Conjectandi* de Jacobi Bernoulli (1655-1705), onde ele conseguiu reunir grande parte dos problemas tratados por seus antecessores de forma sistemática e com demonstrações matemáticas rigorosas. Pode-se citar como exemplo o tratado de probabilidades de Huygens, *De Ratiociniis in Ludo,*[16] cujos conteúdos foram reformulados e generalizados por Bernoulli. As várias correspondências trocadas entre Bernoulli e Leibniz contribuíram para que o cálculo das probabilidades também pudesse ser aplicado em questões civis.[17] Para tratar dessas questões, Bernoulli reservou a quarta parte da sua *Ars Conjectandi* onde contou com a influente contribuição de Leibniz (1646-1716) que tinha formação na área filosófica e jurídica. De acordo com Raymond (1999, p.136), “Leibniz foi um dos primeiros a abrir um caminho para o uso estatístico das probabilidades no campo social, especialmente demográfico.”

---

[14] Huygens, C.: *Oeuvres complètes*, published by the Société Hollandaise des Sciences, 22 vols., The Hague, 1888-1950. Cf. Obra citada em “The Origins of the Infinitesimal Calculus”. Baron, M. Oxford, Pergamon, 1969.

[15] “la composition d'un traité pédagogique, dans l'explicitation de la notion de chance, l'extension méthodologique des solutions de Pascal.”

[16] Bernoulli, 2006, p. IX.

[17] Mais especificamente no capítulo IV da quarta parte da Arte das Conjecturas.

Sylla (1998, pp. 41-76.) também destaca que:

> De 15 de dezembro de 1687, quando ainda trabalhava arduamente em questões relacionadas à Ars Conjectandi, até sua morte em 1705, Bernoulli escreveu repetidamente a Leibniz, pedindo-lhe casos civis, morais e econômicos aos quais ele poderia aplicar sua teoria na Parte 4.

As discussões com Leibniz sobre embate entre probabilidades *a priori* e *a posteriori*, contribuíram para Bernoulli formalizar o teorema fundamental da Ars Conjectandi que posteriormente ficou conhecido como “lei dos grandes números”.[18] A influência dessas questões chega ao século XX, quando Carnap (1963, p. 308, tradução nossa) discute o tema na seguinte passagem:

> Em particular, no que diz respeito aos resultados dos jogos de azar, a ‘probabilidade a priori’ é usada se a evidência fornecer informações apenas sobre as condições gerais do jogo (por exemplo, simetria de um dado ou roleta, semelhança física de cartas e similares), enquanto 'probabilidade a posteriori’**,** refere-se a evidências incluindo resultados estatísticos de jogos anteriores.[19]

Já no que diz respeito a Laplace (1749-1827) encontramos em seus Ensaios Filosóficos de Probabilidades boas evidências de que há acentuada aproximação com pensamento filosófico de Leibniz, em particular, sobre a relação intrínseca entre probabilidade e incerteza. Veremos mais adiante alguns exemplos nos conceitos da teoria das probabilidades de ambos, que confirmam esta aproximação.

Mostraremos também que as questões epistemológicas levantadas por Jaynes, podem ser discutidas através do conceito de indiferença estabelecido por Leibniz (2017, p.162), onde ele afirma que “que há na indiferença uma liberdade onde nada nos obriga a escolher uma ou outra parte.” Em outras palavras, se não há nenhuma razão ou nenhum motivo para escolher um ou outro evento, podemos considerar que todas essas escolhas são igualmente possíveis. Veremos adiante que esse pensamento pode esclarecer, mesmo que parcialmente, os fundamentos epistemológicos do MaxEnt relacionados ao problema da atribuição de probabilidades iniciais.

---

[18] Segundo Jaynes “lei fraca dos grandes números” (Jaynes, 1678, p.213).

[19] In particular, with respect to results of games of chance, 'probability a priori' is used if the evidence gives information only about the general conditions of the game (e.g., symmetry of a die or roulette, physical similarity of cards, and the like), while 'probability a posterior’ refers to evidence including statistical results of earlier games.

Elaboramos esta tese utilizando como principais fontes de consulta livros, artigos e dissertações que estão relacionadas em nossas referências bibliográficas. Através dessas fontes de consulta procuramos revelar os princípios e os métodos que Bernoulli, Laplace e Leibniz aplicaram nas suas estimativas de probabilidades para, em seguida, compararmos esses princípios a fim de reconhecer os possíveis pontos de aproximação entre os conceitos de probabilidades de Bernoulli e Leibniz, Bernoulli e Laplace, Leibniz e Laplace. Lembramos aqui que apenas Bernoulli e Laplace foram citados por Jaynes em seu MaxEnt. No entanto, defenderemos que a filosofia leibniziana tem os fundamentos necessários à compreensão dos vários obstáculos inerentes ao problema da atribuição de probabilidades iniciais.

Nesse panorama, não poderíamos deixar de consultar a obra de Keynes que, em seu Tratado de Probabilidades[20] denominou o controverso princípio da razão insuficiente como princípio da indiferença. De acordo com Keynes, também assumimos que o princípio da indiferença é uma denominação mais apropriada ao problema da atribuição de probabilidades iniciais do que princípio da razão insuficiente, considerando que é possível encontrar uma fundamentação filosófica para o princípio da indiferença no próprio conceito de indiferença em Leibniz. Faremos isso a partir de uma análise em seus textos lógicos e metafísicos e, principalmente, em sua obra *Teodicéia* onde ele aborda enfaticamente os problemas que envolvem a razão, a indiferença e a incerteza dos eventos contingentes.

Sabemos que outros filósofos que antecederam a Leibniz também contribuíram para as questões que pretendemos discutir nesta tese, porém justificamos nossa delimitação nesse filósofo, devido ao alcance do seu pensamento na compreensão das questões epistemológicas apontadas por Jaynes no MaxEnt. Abordaremos essas questões sem a pretensão de fazer uma análise estritamente filosófica, pois também pretendemos ressaltar as possíveis implicações desses conceitos no problema fundamental do MaxEnt que é o da atribuição de probabilidades iniciais.

O objetivo geral desta tese é mostrar que Leibniz foi um dos precursores de conceitos que contribuíram para que Jaynes estabelecesse o MaxEnt. Para alcançar este objetivo, partimos de um estudo preliminar sobre o problema da atribuição de probabilidades, primeiro em Bernoulli e depois em Laplace, para

[20] “A Treatise on Probability" de John Maynard Keynes, (1921, p.46).

confirmar nossas expectativas de que esse problema já se encontrava anteriormente nos manuscritos de Leibniz sobre probabilidades.

De acordo com o que foi exposto até aqui esta tese ficou organizada na forma a seguir.

No primeiro capítulo revisamos o método do MaxEnt destacando as questões epistemológicas que estão contextualizadas no objetivo dessa pesquisa e também um pouco de sua parte histórica. Para isso utilizamos inicialmente como literatura básica o trabalho de Jaynes (1957) onde foi estabelecido o MaxEnt. Também recorremos a outros trabalhos mais específicos que estão relacionados a esse assunto.

No segundo capítulo procuramos confirmar a declaração de Jaynes (1978) de que a base subjacente das atribuições iniciais de probabilidades iniciais constitui um princípio formal na Arte das Conjecturas de Jacobi Bernoulli. Também constatamos através de suas correspondências e bibliografias a importante contribuição de Leibniz para a sua Teoria das Probabilidades, em particular para a quarta parte da *Arte* onde Bernoulli trata da aplicação das probabilidades para as questões civis.

No terceiro capítulo analisamos a obra de Laplace *Théorie analytique des Probabilités* que foi publicada pela primeira vez em 1812 e que deu origem, em 1814, *ao Essai Philosophique sur les Probabilités*. Procuramos identificar os conceitos da teoria laplaciana das probabilidades que podem ter levados outros autores a vincular o seu nome ao denominado princípio da razão insuficiente. Confirmamos nesse capítulo que Laplace não atribuiu essa denominação a nenhum dos dez princípios explicitados em seu Ensaio filosófico sobre as probabilidades.

No quarto capítulo destacamos alguns fatos históricos da formação inicial de Leibniz, como matemático amador, enfatizando a importância dos seus estudos inicias sobre séries e combinações para seus trabalhos posteriores de probabilidades. Confirmou-se esse fato em suas correspondências com Bernoulli e, mais efetivamente, em seus manuscritos de probabilidades sobre os jogos de azar. Desses manuscritos selecionamos os exemplos que incluem o problema da atribuição de probabilidades uniformes e, concomitantemente, os argumentos e os princípios utilizados por Leibniz para justificar esse tipo de probabilidade. Mostramos também que esses argumentos estão muito próximos daqueles utilizados por Bernoulli e Laplace em suas teorias das probabilidades.

Finalmente procuramos ressaltar como o conceito de indiferença encontra-se na filosofia leibniziana como um princípio bem fundamentado. Essa parte da tese constitui o objetivo específico desse capítulo, que é revelar a contribuição de Leibniz no que mais tarde será denominado de princípio da indiferença. No desdobramento desse princípio encontram-se os conceitos de escolha e indiferença de equilíbrio que possibilitam, entre outras questões, discutir o problema da simetria que é um dos principais obstáculos para atribuir probabilidade inicial, como ressaltou Jaynes (1957, p. 622) em seu MaxEnt.

Finalmente reservamos o quinto capítulo para apresentar nossas conclusões finais e também propor algumas perspectivas para novas pesquisas.

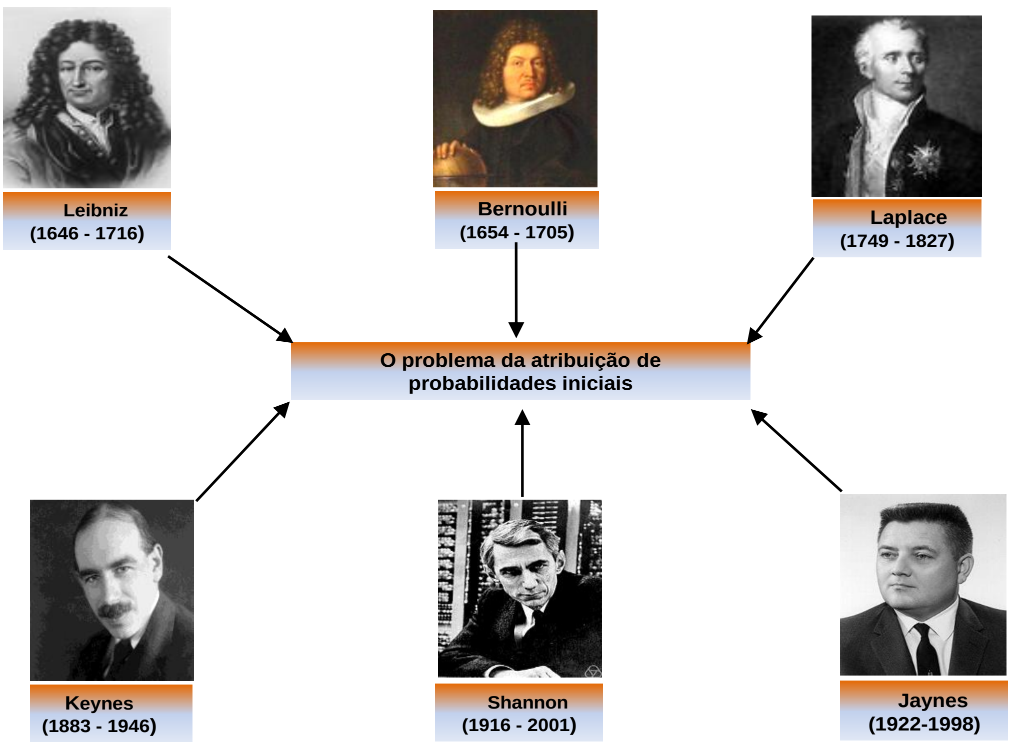


Figura 1: A inclusão de Leibniz entre os precursores dos fundamentos do Princípio da Máxima Entropia. (autoria nossa)

Fonte: Leibniz (www.canstockphoto.com.br/foto-imagens/leibniz.html);
Bernoulli (pt.wikipedia.org/wiki/Jakob_Bernoulli)
Laplace (www.gettyimages.pt/fotos/pierre-simon-laplace)
Keynes (www.gettyimages.com/photos/john-maynard-keyne)
Shannon (en.wikipedia.org/wiki/Claude_Shannon)
Jaynes (en.wikipedia.org/wiki/Edwin_Thompson_Jaynes)
(visitadas em 10/07/2017)

2 O Princípio da Máxima Entropia

No ano de 1948 o engenheiro norte americano Claude E. Shannon publicou a *Teoria Matemática da Informação* onde se estabelece que "o problema fundamental da comunicação é reproduzir em um ponto exatamente ou aproximadamente uma mensagem selecionada em outro ponto." (SHANNON, 1948, P.1). São vários obstáculos que podem dificultar a comunicação entre dois pontos distintos que podem começar nas fontes de informação[21] até os respectivos receptores. Essas fontes, como relés, telégrafos, câmeras de televisão, rádio e telefonia, possuem uma capacidade ou uma taxa de transmissão associada e que pode ser medida em bits por segundo. Nessas condições, a informação só pode ser transmitida através do canal se a quantidade de informação enviada pela fonte do canal não exceder a sua capacidade de transmissão. Por esse motivo e outros subjacentes, Shannon procurou otimizar com sua teoria da informação os meios de comunicação daquela época, diagnosticando a capacidade de transmissão de cada canal e o nível de confiabilidade das informações enviadas, desde o emissor até o receptor.

Possivelmente a escassez de uma tecnologia mais confiável disponível no meado do século XX e a limitação dos meios de comunicação daquela época foram alguns dos fatores que conduziram Shannon a tratar de problemas relacionados aos ruídos que ocorriam nos canais durante a transmissão de informações[22]. Diante desse problema "a correção adequada a ser aplicada à quantidade de informação transmitida é a quantidade dessa informação que está faltando no sinal recebido ou na incerteza quando recebemos um sinal do que realmente foi enviado." (SHANNON, 1948, p.20)

A questão que se coloca diante dos obstáculos destacados acima é procurar um meio para quantificar a informação transmitida pelo respectivo canal de comunicação. Nesse caso a expressão H = $-\sum p_i ln p_i$ desempenha um papel fundamental na teoria da informação (SHANNON, 1948, p.11). Ela fornece as

---

[21] Entre essas fontes Shannon (1948, p. 2) destaca: (a) Uma sequência de letras como em um telégrafo do sistema de teletipo; (b) Uma única função de tempo f(t) como em rádio ou telefonia; (c) Uma função de tempo e outras variáveis como na televisão em preto e branco - aqui a mensagem pode ser pensada como um função f (x; y; t) de duas coordenadas espaciais e tempo, a intensidade da luz no ponto (x; y) e o tempo t em um placa de tubo de captação; (d) Duas ou mais funções do tempo, digamos f (t), g(t), h(t) - este é o caso da transmissão de som "tridimensional" ou se o sistema se destina a atender vários canais individuais em multiplex; (e) Várias funções de várias variáveis - na televisão a cores a mensagem consiste em três funções f (x; y; t), g(x; y; t), h(x; y; t) definidas em um continuum tridimensional.

[22] O canal é apenas o meio usado para transmitir o sinal do transmissor para o receptor. Pode ser um par de fios, um cabo coaxial, uma banda de radiofrequências, um feixe de luz, etc. (SHANNON, 1948, p.2)

medidas de informação e da respectiva incerteza associadas a uma certa mensagem, cuja probabilidade é dada por p_i. Pode-se observar que H tem a mesma expressão da entropia da termodinâmica $-K \sum p_i ln p_i$, exceto pela presença da constante K de Boltzmann. Porém, é preciso ter cautela quanto a essa correspondência entre a expressão da entropia na teoria da comunicação e da entropia na termodinâmica, ou seja:

> O mero fato de que a mesma expressão matemática $-\sum p_i \ln p_i$ ocorrer 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* [...]. O recurso que estava faltando foi fornecido recentemente por Shannon na demonstração de que a expressão da entropia tem um significado mais profundo, independente da termodinâmica. (JAYNES, 1957, 621, grifo do autor, tradução nossa)[23]

Conforme já ressaltamos na introdução, o conceito de entropia, que anteriormente estava vinculado somente aos problemas da física, pôde pelo MaxEnt ser aplicado a uma grande variedade de problemas causando um grande impacto em diversas áreas do conhecimento científico. 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. Segundo Jaynes (1957, p. 620, tradução nossa)[24] "A teoria da informação forneceu um critério construtivo para atribuir distribuições de probabilidade com base no conhecimento parcial e leva a um tipo de inferência estatística que é chamada estimativa de entropia máxima".

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:

---

[23] The mere fact that the same mathematical expression $-\sum p_i \ln p_i$ occurs both in statistical mechanics and in information theory does not in itself establish any connection between these fields. This can be done only by finding new viewpoints from which thermodynamic entropy and information-theory entropy appear as the same concept [...].The feature which was missing has been supplied only recently by Shannon in the demonstration that the expression for entropy has a deeper meaning, quite independent of thermodynamics.

[24] "Information theory provides a constructive criterion for setting up probability distributions on the basis of partial knowledge, and leads to a type of statistical inference which is called the maximum-entropy estimate."

-De quantas maneiras diferentes um determinado número de moléculas pode ser distribuído?

- Entre todas as distribuições possíveis, qual é a mais provável?

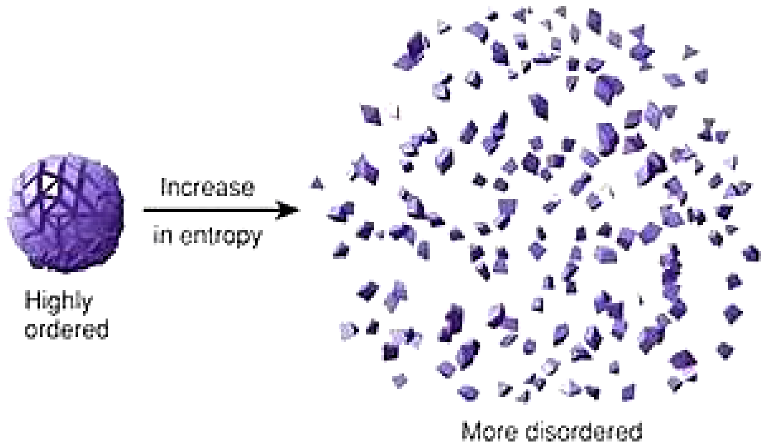


Figura 2: Ilustração que relaciona entropia e desordem.
Fonte: https://tse1.mm.bing.net/th?id=OIP.qkiOvV7gNO6qmXZRZA7BagHaEM&pid=Api&P=0&w=304&h=173- (visitado em 12/07/2017)

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 restrições (JAYNES, 1978, p.15). Seguindo essa linha de raciocínio, podemos dizer que para fazer inferências com base em informações incompletas ou quando não houver nenhuma informação, devemos usar a distribuição de probabilidade que tenha a máxima entropia. Nesse caso, a distribuição uniforme é a mais imparcial ou menos tendenciosa possível e reflete o desconhecimento inicial e a incerteza que se tem quanto ao estado de um sistema.

Nessas condições, o problema a ser resolvido era obter um método que não fosse tendencioso e que estivesse de acordo com as informações disponíveis. Boltzmann procurou resolvê-lo direcionando seus esforços na interpretação estatística da termodinâmica. Shannon procurou concentrar-se na otimização das linhas de comunicação. Com o MaxEnt, Jaynes (1957, p. 623) procurou unificar essas duas áreas científicas quando sugere que a mecânica estatística pode ser independentemente de qualquer argumento físico e, em particular, independentemente de qualquer verificação experimental.

Conforme já foi mencionado anteriormente, a expressão matemática $-\sum p_i ln p_i$ é a mesma que ocorre tanto na mecânica estatística como na teoria da informação. Na primeira ela aparece como a entropia termodinâmica e na outra como entropia da teoria da informação. Em ambos os casos, a quantificação da incerteza de um sistema, através de uma distribuição discreta de probabilidade, torna-se um problema central.[25] Essa incerteza é decorrente da falta de informação ou de informações incompletas e, para diferentes quantidades de informações, teremos associadas diferentes distribuições de probabilidade. Desse modo, quanto maior for a quantidade de informações menor será a entropia do sistema. Por outro lado, quando não houver nenhuma informação disponível a incerteza ou a entropia será máxima e o sistema é o mais aleatório possível. Segundo Jaynes, “uma ampla distribuição de probabilidade representa mais incerteza do que uma precisão acentuada”.[26] Portanto, o conceito de entropia está diretamente associado à incerteza tanto na teoria da informação quanto no MaxEnt.[27]

O Princípio da Máxima Entropia estabelece uma forma de tomar a entropia como um conceito de partida, conforme veremos na sua formalização matemática. Apresentaremos aqui apenas um resumo já que há uma vasta bibliografia (já citada) e que trata desse assunto detalhadamente.

Supomos que certa quantidade x pode ter valores discretos $x_i \ (i = 1, \ldots, n)$, mas não sabemos a probabilidade correspondente p_i. Tudo que sabemos é o valor esperado de uma certa função do sistema *f* (*x*),

$$< f(x) > = \sum_{i=1}^{n} p_i \, f(x_i) \qquad (\mathbf{2.1})$$

A partir dessa informação como poderemos obter o valor esperado de outra função do sistema *g* (*x*)? A informação dada é insuficiente para responder a este problema, ou seja, para se determinar a probabilidade p_i. A equação (2.1) com a condição de normalização

$$\sum_{i=1}^{n} p_i = 1 \qquad (\mathbf{2.2})$$

[25] "amount of uncertainty" represented by a discrete probability distribution. (Jaynes, 1957, p.621)
[26] Jaynes, E. T. ``Prior Probability", IEEE Transations On Systems Science and Cybernetics, vol.4 sec.4 N.3 (1968).
[27] Jaynes (1957, p. 622) considerou os termos “entropia” e “incerteza” como sinônimos.

não são suficientes para se resolver o sistema de equações e encontrar p_i. As equações (2.1) e (2.2) são denominadas vínculos do sistema. Adiciona-se a estes dois vínculos a expressão de uma certa quantidade H (p_i ... p_i), denominada entropia da distribuição de probabilidade p_i, que é a mesma da Teoria da Informação de Shannon (1948).

$$H(p_1 \dots . p_i) = -K \sum_{i=1}^{n} p_i ln p_i \qquad (\mathbf{2.3})$$

A expressão (2.3) é a mesma expressão da Mecânica Estatística e é chamada de "*entropia da distribuição de probabilidade* p_i", onde K uma constante positiva. Para maximizar *H*, sujeita aos vínculos (2.1) e (2.2), basta extremizar a função (2.3) pelo método tradicional dos multiplicadores de Lagrange, que é utilizado em várias áreas da Física e de Matemática. Aqui denotaremos estes dois multiplicadores respectivamente por *λ e μ,* associados aos vínculos (2.1) e (2.2).

A função a ser maximizada é

$$\phi = -K \sum_{i=1}^{n} f_i ln f_i \; + \; \lambda \{ \sum p_i \, f(x_i) \; -< f(x) >\} + \mu \{ \sum (p_i - 1) \} \qquad (\mathbf{2.4})$$

e assim obteremos a distribuição de probabilidade

$$p_i = \; e^{-\mu - \lambda f(x)} \qquad (\mathbf{2.5})$$

a qual maximiza a entropia (2.3) sujeita aos vínculos (2.1) e (2.2).

Há um fato interessante para o caso de não termos nenhuma informação do sistema, como a (2.1), que é o valor médio de algum observável do sistema. Neste caso nos resta apenas o vínculo da condição de normalização (2.2). Pela consistência do MaxEnt, isto nos deverá levar a uma distribuição de probabilidade uniforme. Veja-se que se não há valor esperado da quantidade < f(x_i) >, não há o termo correspondente em (2.4), e em (2.5) a probabilidade se reduz a $p = \; e^{-\mu}$ que ao ser substituída em (2.2) nos fornece

$$\sum_{i=1}^{n} e^{-\mu} = \; 1 \qquad (\mathbf{2.6})$$

que equivale a

$$e^{-\mu} + \ e^{-\mu} + \ e^{-\mu} + \dots + e^{-\mu} = 1.$$

Daí temos

$$n\, e^{-\mu} = 1 \quad \Rightarrow \qquad e^{-u} = \frac{1}{n} \ , \qquad (2.7)$$

que é exatamente a distribuição de probabilidade uniforme, conforme esperávamos. Sendo assim, o MaxEnt também vale para o caso onde não há nenhuma informação sobre o sistema. É nesse sentido que compreendemos a afirmação de Jaynes (1957, p.623) de que "o princípio da máxima entropia pode ser considerado como uma extensão do princípio da razão insuficiente", pelo qual se estabelece que todas as probabilidades devem ser iguais se não houver nenhuma razão para se pensar ao contrário. Desse modo, o conceito de entropia forneceu o critério de escolha que faltou a Laplace necessário para eliminar a aparente arbitrariedade[28] do princípio da razão insuficiente, que tinha como único argumento a falta de uma razão para atribuir probabilidades diferentes a vários eventos possíveis.

Entretanto, cabe-nos perguntar, de acordo com Leibniz (1974, p.273), "Como determinar com precisão os limites que demarcam a razão?" Sem a pretensão de responder essa pergunta integralmente, podemos dizer que é questionável falar em "razão insuficiente" sem envolvimento de juízos subjetivos, pois essa condição reflete um estado de conhecimento que, no mínimo, poderíamos classificá-lo como vago ou incompleto. Além disso, o que é razoável para um determinado ponto de vista pode não ser para outro. Por exemplo: uma pessoa vai assistir uma corrida de carros esportivos e desconhece a capacidade de cada piloto e a potência de cada veículo. Ela sabe apenas que 25 carros irão competir. Nessas condições, ela diz que todos os carros têm a mesma probabilidade de vencer, pois não há nenhuma diferença ou nenhuma razão para que um deles ganhe em vez de outro qualquer.

[28] Jaynes (1957, p. 623)

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.

Nesse sentido podemos dizer junto com Jaynes (1957, p.622, tradução nossa) [29] que

> a escola do pensamento "subjetivo" considera as probabilidades como expressões da ignorância humana; a probabilidade de um evento é meramente uma expressão formal de nossa expectativa de que um evento acorreu ou ocorrerá, com base em alguma informação que esteja disponível.

Sendo assim, as estimativas iniciais de probabilidades dependem das informações que estão disponíveis, onde a situação mais incerta é aquela em que não temos nenhuma informação, o que leva à distribuição de probabilidades uniformes. Naturalmente, quanto maior for o número de informações, menor será nossa incerteza e, consequentemente, as distribuições de probabilidades serão cada vez mais precisas. Suponhamos, por exemplo, que alguém nos diga que uma caixa contém seis bolas com as cores azul, verde e preta. Porém, não sabemos quantas bolas de cada cor foram colocadas na caixa. Pergunta-se: qual a probabilidade de se retirar uma bola verde? Como não sabemos quantas bolas verdes foram colocadas na caixa, afirmamos, por questão de coerência, que há uma probabilidade de 1/6 para retirá-la, pois de acordo com as informações disponíveis, não há nenhuma razão para dizer que há mais probabilidade de se retirar uma bola verde em vez de outra qualquer. Além disso, a única garantia que temos é de que há pelo menos uma bola verde dentro da caixa. Mas se recebemos a informação de que dentro da caixa tem duas bolas azuis e uma preta, podemos afirmar que a probabilidade de retirar uma bola azul é 2/6, uma bola verde 3/6 e uma bola preta 1/6. Isso nos mostra que quanto maior for número de informações menor será a incerteza e as probabilidades tornar-se-ão gradativamente mais acentuadas.

Concisamente, podemos dizer de que o MaxEnt possibilita tratar vários fenômenos que envolvem imprevisibilidade e incerteza, como ocorre frequentemente na análise de vários problemas, conforme nos referimos na Introdução. Para começarmos a dirigir nossa abordagem para o nosso ponto

---

[29] the ‘subjective’ school of thought regards probabilities as expressions of human ignorance; the probability of an event is merely a formal expression of our expectation that the event will or did occur, based on whatever information is available.

principal, que é o princípio da indiferença, vejamos como Jaynes (1957, p.626, grifo do autor, tradução nossa) se refere a essas questões.

> A entropia como conceito pode ser considerada uma medida do nosso grau de ignorância quanto ao estado de um sistema; por outro lado a entropia é, para as condições de equilíbrio, uma quantidade experimentalmente mensurável, cujas propriedades mais importantes foram encontradas empiricamente. [...] Alguém pode então perguntar como essas probabilidades poderiam ser de alguma forma relevantes para o comportamento de sistemas físicos reais. A boa resposta a esta questão é a famosa observação de Laplace que a teoria da probabilidade não é nada senão o senso comum reduzido ao cálculo'. Se tivermos pouca ou nenhuma informação relevante para uma certa questão, o senso comum nos diz que nenhuma conclusão pode ser fortemente justificada. [30]

Complementando essa passagem, podemos avançar um pouco com seguinte pensamento de Leibniz (1999, p.689, tradução nossa): [31]

> Os graus de probabilidade que existem nas conjecturas podem fornecer uma estimativa tão segura quanto os números. Entretanto, essa estimativa não pode e não deve ser usada para chegar a uma certeza, o que é impossível, mas para agir de maneira mais razoável possível sobre os fatos ou conhecimentos dados a nós.

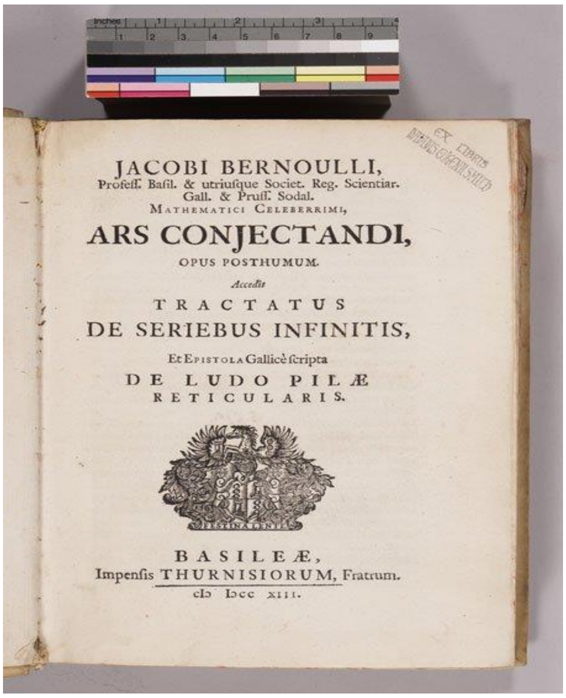
Concluímos, a partir das questões abordadas acima, que o emprego de probabilidades para descrever uma determinada situação envolve, necessariamente, a incerteza. Sendo assim, cada distribuição de probabilidade distingue-se uma das outras pelo seu grau de incerteza que, como já destacamos anteriormente, é máxima quando essa distribuição é uniforme. Além disso, diferentes graus de incerteza estão associados a diferentes distribuições de probabilidades.

[30] Entropy as a conconcept may be regarded as a measure of our degree of ignorance as to the state of a system; on the other hand, for equilibrium conditions it is an experimentally measurable quantity, whose most important properties were first found empirically. [...] One might then ask how such probabilities could be in any way relevant to the behavior of actual physical systems. A good answer to this is Laplace's famous remark that probability theory is nothing but 'common sense reduced to calculation.' If we have little or no information relevant to a certain question, common sense tells us that no strong conclusions either way are justified."

[31] les degrés de probabilité qu'il y a dans les conjectures qui ont leur estimation aussi asseurée que les nombres; cette estimation nous peut et doit servir non pas pour venir à une certitude, ce qui est impossible mais pour agir le plus raisonnablement qu'il se peut sur les faits ou connoissances qui nous sont données.

3 JACOBI BERNOULLI

As várias biografias e os estudos sobre a obra e a vida de Bernoulli apontam a Arte das Conjecturas[32] (doravante *Arte*) como um dos mais importantes trabalhos deste matemático. Nessa obra encontram-se os princípios de sua teoria das probabilidades com suas respectivas demonstrações matemáticas. Entre esses princípios é do nosso interesse destacar e analisar aqueles que têm alguma relação com as estimativas de probabilidades iniciais. Porém, antes de iniciarmos nossa abordagem a esse problema destacaremos algumas características e alguns fatos históricos que precederam a construção da *Arte* e também a influência de Leibniz nessa obra. Embora esses fatos já sejam bem conhecidos, optamos por expô-los aqui simplesmente como um critério pedagógico para que essas informações estejam também ao alcance daqueles que não são familiarizados com a obra desse autor. Todavia, não se trata de um resumo da obra e nem da biografia de Bernoulli.



JACOBI BERNOULLI,
Profess. Basil. & utriusque Societ. Reg. Scientiar.
Gall. & Pruss. Sodal.
MATHEMATICI CELEBERRIMI,
ARS CONJECTANDI,
OPUS POSTHUMUM.
Accedit
TRACTATUS
DE SERIEBUS INFINITIS,
Et EPISTOLA Gallicè scripta
DE LUDO PILÆ
RETICULARIS.
FESTINA LENTE
BASILEÆ,
Impensis THURNISIORUM, Fratrum.
cIↄ Iↄcc XIII.

Figura 3: Folha de rosto da Ars Conjectandi de Jacobi Bernoulli, publicada postumamente em 1713 por seu sobrinho Nicolaus I Bernoulli.
Fonte: https://www.maa.org/press/periodicals/convergence/mathematical-treasures-jacob-bernoullis-ars-conjectandi. (visitado em 2/12/2019)

[32] "A arte da conjectura, ou estocástica, é a arte de medir as probabilidades das coisas da forma mais exata possível." (Bernoulli, 2006, pp. 317-318)

Jacobi Bernoulli nasceu na Suíça em 27 de dezembro de 1654. Estudou teologia por insistência do seu pai, pois desde jovem manifestava extraordinária vocação para a matemática. Fez várias viagens pela Europa onde conviveu com alguns cientistas importantes do século XVII como Jean Hudde, Robert Boyle, Robert Hooke, Edward Stilling e outros. Após o encontro com estes cientistas, Jacobi voltou a dar aulas em Basiléia (Basel), ministrando a disciplina de matemática. A partir de 1680, fez várias publicações com assuntos científicos diversificados como o estudo dos cometas, séries harmônicas, o problema da braquistócrona, equações diferenciais e, sobretudo, o cálculo infinitesimal leibniziano (Bos, 1985, p.4). Em 1682, dedicou-se a leitura dos trabalhos de Leibniz, que naquela ocasião era o supervisor da Acta Eruditorum (Ata dos Eruditos).

Logo após Leibniz publicar seu primeiro trabalho sobre o Cálculo Diferencial no Acta Eruditorum de Leipzig, em outubro de 1684, Jacobi e seu irmão Johan adotaram esse novo método, aplicando-o a vários problemas de matemática que publicaram em várias revistas científicas. Assim começou a dinastia Bernoulli de matemáticos que duraria várias gerações. Jacobi e Johann I foram logo seguidos por seu sobrinho Nicolaus I. Os Bernoullis viajaram para países estrangeiros para assumir posições em matemática, mas preferiram trabalhar em Basileia sempre que possível. Nicolaus Bernoulli I assumiu a cadeira de matemática em Pádua, sucedendo a Jacobi Hermann em 1716.

Em suas viagens pela Europa, Leibniz incentivou a criação de vagas para professores de matemática em nível superior nas universidades. Com isso, Jacobi Bernoulli foi eleito membro estrangeiro da Academia de Ciências Parisiense e da Academia de Ciências de Berlim. Johann Bernoulli I tornou-se membro da Academia de Paris em 1699, da Academia de Berlim em 1701 e da Royal Society de Londres em 1712. Logo depois, em 1714, Nicolas também se tornou membro da Royal Society. Assim, os matemáticos da família Bernoulli deixaram seus nomes marcados no cenário científico internacional.

Na construção da sua teoria das probabilidades, Jacobi Bernoulli pretendia aplicar o cálculo das probabilidades além das fronteiras dos jogos de azar, por exemplo, nas questões jurídicas e econômicas. Como ele não tinha formação na área jurídica ele escreveu repetidamente a Leibniz, pedindo-lhe algumas questões sobre esses assuntos para que pudesse incluí-los na Parte IV do seu trabalho. Nas cartas (SYLLA, 1968) enviadas a Bernoulli, Leibniz enfatizou a importância da estimativa de probabilidades não só no tratamento matemático dos vários tipos de jogos, mas também nas questões sociais. A partir dessas correspondências várias questões sobre a estimativas de probabilidades foram discutidas por ambos na elaboração final da *Arte*. Algumas dessas questões estão incluídas no objetivo desta tese como veremos no decorrer deste capítulo.

A primeira parte da *Arte* é basicamente uma reformulação do trabalho de Huygens sobre jogos, intercalada com algumas observações, generalizações e outros métodos alternativos. A segunda parte contém uma exposição sistemática da matemática de combinações e permutações baseada em alguns trabalhos anteriores. Como o próprio Bernoulli (2006, p.193, tradução nossa) declarou, “vários homens ilustres, van Schooten, Leibniz, Wallis e Prestet, optaram por abordar esse assunto, para que ninguém suponha que tudo o que estamos prestes a dizer aqui seja novo.” [33]

A terceira parte aborda as aplicações das combinações e permutações aos vários tipos de jogos e nela Bernoulli (2006, p.251, tradução nossa) estabeleceu o seguinte princípio. “O fundamento geral consiste em tomar todas as combinações e permutações como casos equipossíveis e em considerar diligentemente quantos desses casos são favoráveis ou opostos a este ou àquele jogador.” [34] Nesse princípio encontra-se subentendida a ideia de uma equiprobabilidade (casos equipossíveis) e também a própria definição de probabilidade como a relação entre o número de casos favoráveis e todos os casos possíveis que são constituídos por combinações e permutações. Como ressaltou Jaynes (1978, p.213, tradução nossa) [35], “essa é a única crítica válida que se pode fazer a esse princípio pois, em sua forma original, a enumeração dos casos igualmente possíveis não pode ser aplicada a todos os tipos de problemas.” De fato, podemos confirmar essa afirmação de Jaynes nas próprias palavras de Bernoulli (2006, p.326, tradução nossa).

> a única coisa necessária para formar conjecturas corretamente sobre algum assunto é determinar o número desses casos com precisão e determinar qual deles pode acontecer com mais facilidade do que os outros. Mas aqui chegamos a um impasse, pois

---

[33] “several distinguished men, namely van Schooten, Leibniz, Wallis, and Prestet, have chosen to take up this matter for treatment lest anyone assume that all of what we are about to say here is new.”
[34] “The general foundation consists in taking all the combinations and permutations [...] as so many equipossible cases and in diligently considering how many of these cases are favorable to or opposed to this or that player.”
[35] “The only, valid criticism of this principle, it seems to me, is that in the original form enumeration of the equally possible cases it cannot be applied to all problems.”

> isso dificilmente pode ser feito. Na verdade, dificilmente pode ser feito em qualquer lugar, exceto em jogos de azar. Os inventores desses tipos de jogos se esforçaram para torná-los justos de tal forma que todos os resultados sejam conhecidos e aconteçam com igual facilidade. Mas isso de modo algum ocorre com a maioria dos outros eventos que dependem da operação da natureza ou da vontade humana. [36]

Como vimos na passagem acima, a atribuição de probabilidades iniciais ou probabilidades a priori, como se refere Bernoulli em sua Arte das Conjecturas, só se aplica praticamente aos jogos de azar, como nos mostra o seguinte exemplo.

> Em um dado todos os resultados têm *chances iguais* para ocorrer; por causa da semelhança de suas faces e do seu peso uniforme, sendo assim *não há razão* para que uma das faces seja mais propensa a cair do que outra. Isso também ocorre quando se conhece o número de fichas brancas ou pretas de uma urna. Sabe-se que todos esses casos são *igualmente possíveis***,** pois *não há razão* para que uma ficha seja retirada em vez do outra. [...] Mas quem tem uma perspectiva suficiente sobre a natureza da mente humana quem ousaria determinar os casos em que este ou aquele jogador pode ganhar ou perder no jogo? Seria desejável que alguém aprendesse qualquer coisa sobre essas e outras situações semelhantes, uma vez que são altamente dependentes de uma variedade incontável de combinações. (BERNOULLI, 2006, p.326, grifo nosso, tradução nossa). [37]

Fica claro no exemplo acima que a ideia de um dado equilibrado e uniforme é a condição que serve como argumento para atribuir chances iguais a todas as seis faces do dado. No caso das fichas essa condição está na equivalência entre o número de fichas brancas e pretas. Nas duas situações vale o princípio da indiferença, pois “não há nenhuma razão” para privilegiar um ou outro resultado qualquer. De outro modo podemos dizer que dois ou mais eventos são “igualmente possíveis” se, e somente se, não houver nenhuma razão para se afirmar ao contrário. Na maioria das vezes, relaciona-se este enunciado ao princípio da razão

---

[36] From this it resulted that the only thing needed for correctly forming conjectures on any matter is to determine the numbers of these cases accurately and then to determine how much more easily some can happen than others. But here we come to a halt, for this can hardly ever be done. Indeed, it can hardly be done anywhere except in games of chance. The originators of these games took pains to make them equitable by arranging that the numbers of cases resulting in profit or loss be definite and known and that all the cases happen equally easily. But this by no means takes place with most other effects that depend on the operation of nature or on human Will.

[37] So, for example, the numbers of cases in dice are known. Moreover these all have equal tendencies to occur; because of the similarity of the faces and the uniform weight of the die, there is no reason why one of the faces should be more prone to fall than another. In the same way the numbers of cases for drawing white or black slips of paper from an urn are known. It is also known that they are all equally possible, because, without doubt, the number of slips of each type is known and determined and there is no reason why one of them should be drawn from the urn rather than another. [...] Again, who has a sufficient perspective on the nature of the human mind or on the wonderful structure of the body so that they would dare to determine the cases in which this or that player may win or lose in games that depend in whole or in part on the shrewdness or the agility of the players? In these and similar situations, since they may depend on causes that are entirely hidden and that would forever mock our diligence by an innumerable variety of combinations.

insuficiente, mas aqui adotamos a denominação princípio da indiferença que foi aquela sugerida por Keynes, como já foi mencionado anteriormente.

Sabendo a priori que todos os casos são igualmente possíveis, pode-se agora calcular a probabilidade de um evento. Para isso basta saber, entre todos os casos possíveis o número de casos que são favoráveis a esse evento. Assim, sua probabilidade será dada pela razão entre o número de casos que lhe são favoráveis e o número de todos os casos possíveis o que leva, segundo Jaynes (1978, p. 213), à "definição clássica de probabilidade, cuja regra geral é dada por
p (A) = M / N, onde M é o número de casos favoráveis ao evento A e N o número total de casos igualmente possíveis." Entretanto, se o número de casos favoráveis for o mesmo para cada um dos eventos possíveis, então suas probabilidades serão iguais o que leva a probabilidade uniforme (1/n) como caso particular da definição de probabilidade.

Vamos ilustrar esses dois tipos de probabilidades com o seguinte exemplo. Uma urna contém uma ficha azul, uma ficha branca e uma ficha verde. Podemos supor que é igualmente possível retirar qualquer uma dessas fichas, pois nenhuma razão indica o contrário. Além disso, o número casos favoráveis é o mesmo para cada ficha, ou seja, um caso favorável. Sabendo que o número de todos os casos possíveis é três, a probabilidade de retirar qualquer uma dessas fichas é a mesma, isto é, $p(A) = p(B) = p(V) = \frac{1}{3}$. Observamos também que se o número de fichas for aumentado proporcionalmente, em cada um desses casos, ainda assim teríamos a mesma probabilidade para retirar qualquer uma dessas fichas, por exemplo: 2 fichas azuis, 2 fichas brancas e 2 fichas verdes. Nesse caso, o número de casos favoráveis é de duas fichas para cada cor e o número de todos os casos possíveis é 6. Apesar de ter sido aumentado o número de casos favoráveis a probabilidade de se retirar uma dessa fichas, ainda permanece igual ao caso anterior, isto é:

$$p(A) = p(B) = p(V) = \frac{2}{6} = \frac{1}{3},$$

Suponhamos agora que uma urna contém 3 fichas azuis, 2 fichas brancas e 1 ficha verde. Nessas condições a probabilidade de se retirar uma ficha azul é dada pela razão entre o seu número de casos que lhe são favoráveis pelo número de todos os casos possíveis, ou seja, p(A) = $\frac{3}{6}$. Do mesmo modo pode-se encontrar a probabilidade para retirar uma ficha branca que é p (B) = $\frac{2}{6}$ e probabilidade de retirar uma ficha verde que é p(V) = $\frac{1}{6}$. Note que também podemos calcular essas probabilidades multiplicando os respectivos números de casos favoráveis pela probabilidade individual de cada ficha que é $\frac{1}{6}$. Sendo assim temos, p(A) = 3 $\frac{1}{6}$, p(B) = 2 $\frac{1}{6}$, e p(V) = 1 $\frac{1}{6}$. Assim, o que faz distinguir se um evento é mais provável do que outro são seus respectivos "pesos"[38], ou seja, os seus casos favoráveis. Vimos com estes exemplos que as estimativas de probabilidades iniciais são dadas pela razão entre o número de casos favoráveis pelo número de todos os casos possíveis. Mas esse procedimento não se aplica aos eventos que dependem total ou parcialmente da vontade humana, da agilidade ou do raciocínio, como é o caso do jogo de xadrez. Também não se aplica aos eventos que dependem da operação da natureza como o clima e algumas catástrofes. Alguns fenômenos da natureza envolvem tantas circunstâncias e um grande número de causas agindo simultaneamente que se torna inviável estimar probabilidades antecipadamente, pois, nesses casos, não há meios que possibilitem diagnosticar o número de casos que podem ou não favorecer a sua ocorrência. Sendo assim, somos conduzidos a buscar nos resultados anteriores alguma regularidade para fazermos previsões sobre os eventos futuros. Logo, a probabilidade dada a posteriori depende de um grande número de experiências ou de observações sucessivas e, à medida que o evento observado se repete, sua probabilidade aumenta proporcionalmente. Se durante vários anos observou-se que a temperatura média de uma cidade durante o inverno é 10°C, espera-se que no ano seguinte a temperatura esteja em torno dessa medida. Entretanto, isso não é uma verdade absoluta. O problema é que nesse tipo evento e também em outros, as informações anteriores nem sempre servirão como parâmetro para as posteriores. Não é pelo fato de ter nascido mais meninos no ano passado que isso se repetir no presente ou no futuro. Sendo

---

[38] Bernoulli, Arte das conjecturas, p. 321.

assim, seria preciso um período maior de observações para encontrar uma suposta regularidade nos resultados.

Conforme já mencionamos anteriormente, atribuir probabilidades a priori a um determinado evento é preciso contar o número de casos que lhe são favoráveis e o número de todos os casos possíveis para, em seguida, determinar a razão entre eles. Mas esse princípio, segundo Bernoulli, só se aplica praticamente aos jogos de azar. Para eventos imprevisíveis (também chamados de contingentes) ou que dependem de causas totalmente ocultas, Bernoulli estabeleceu o seguinte princípio.

> O que não pode ser averiguado a priori pode, pelo menos, ser descoberto a posteriori a partir dos resultados muitas vezes observados em situações semelhantes, uma vez que se deve presumir que algo que aconteceu no passado pode acontecer ou não no futuro em circunstâncias semelhantes. [...] Da mesma forma, se alguém durante vários anos tivesse observado o clima e anotado quantas vezes estava claro ou chuvoso ou se alguém tivesse assistido com muita frequência dois jogadores em um jogo e anotado quantas vezes esse ou aquele jogador venceu, teria descoberto a razão que provavelmente existe entre o número de casos em que os mesmos resultados podem ocorrer ou não no futuro em circunstâncias semelhantes às anteriores. (BERNOULLI, 2006, p. 327, tradução nossa) [39]

Sobre esse princípio, Leibniz (1995, p.29, tradução nossa) fez a seguinte ressalva.

> Parece-me que a dificuldade para estimar probabilidades empiricamente por resultados sucessivos, está nas coisas contingentes, ou seja, que dependem de uma infinidade de circunstâncias que não podem ser determinadas, pois a natureza tem seus hábitos. Quem pode dizer que o resultado seguinte não irá divergir um pouco da lei de todos os anteriores devido à mutabilidade das coisas? Novas doenças atacam a humanidade e, mesmo que você tenha observado os resultados para qualquer número de mortes, ainda assim não seria possível estabelecer limites para a natureza das coisas, para que elas não pudessem variar no futuro.[40]

Em uma das suas respostas as objeções de Leibniz, Bernoulli respondeu.

> Se as doenças se multiplicarem com o passar do tempo, seria

[39] What cannot be ascertained a priori, may at least be found out a posteriori from the results many times observed in similar situations, since it should be presumed that something can happen or not happen in the future in as many cases as it was observed to happen or not to happen in similar circumstances in the past.[...] Likewise if someone for several years past should have observed the weather and noted how many times it was clear or rainy or if someone should have very frequently watched two players at a game and shotfid have seen how many times this or that player won, just by doing so one would have discovered the ratio that probably exists between the numbers of cases in which the same outcomes can happen or not happen in the future in circumstances similar to the previous ones.

[40] The difficulty in it seems to me to be that contingent things or things that depend on infinitely many circumstances cannot be determined by finitely many results, for nature has its habits, following from the return of causes, but only for the most part. Who is to say that the following result will not diverge somewhat from the law of all the preceding ones because of the mutabilities of things? New diseases attack humankind. Therefore even if you have observed the results for any number of deaths, you have not therefore set limits on the nature of things so that they could nor vary in the future."

> necessário fazer novas observações. É certo que alguém que quisesse julgar a vida daqueles que viveram antes do dilúvio por observações feitas hoje em Londres ou Paris ou em qualquer lugar se afastaria totalmente da verdade? (BERNOULLI, 2006, p.41, tradução nossa) [41]

Essa resposta sugere uma atualização constante de informações e, a partir dessas atualizações, podem-se inferir novos resultados. Trata-se então de um procedimento estatístico cujas probabilidades são obtidas a posteriori, isto é, por resultados sucessivos e oriundos de alguma experiência. Entretanto, esses resultados dependem de frequentes atualizações que, por sua vez, também estarão sujeitas a novas atualizações. Nesse caso, cabe perguntar se os resultados obtidos por meio de experiências sucessivas representam as melhores estimativas que poderiam ter sido feitas com base nas Informações disponíveis?

Podemos encontrar uma possível resposta para essa pergunta em uma carta que Leibniz escreveu para Louis Bourguet (1678-1742) em 1714.

> A arte de conjecturar baseia-se no que é mais ou menos fácil ou mais ou menos factível, porque do latim facilis deriva um faciendo que significa factível. [...] O Sr. Bernoulli cultivou este assunto em minhas exortações. Assim, estimamos as probabilidades a posteriori, *pela experiência*, e devemos recorrer a elas na ausência de razões a priori: por exemplo, é igualmente provável que uma criança que está pra nascer seja um menino ou uma menina, porque o número de meninos e de meninas são quase iguais neste mundo. Podemos dizer que aquilo que acontece mais ou menos vezes é também o mais ou menos factível no estado atual das coisas, combinando todas as considerações que devem contribuir para a produção de um fato. (Leibniz, 1978, p. 570, grifo nosso)

Nessa passagem fica clara participação de Leibniz na construção da teoria das probabilidades de Bernoulli. Ela também nos mostra que o embate entre probabilidades a priori e a posteriori envolveu outros matemáticos como foi o caso de Bourguet. As discussões em torno desse assunto permanecem até hoje, como é possível confirmar nos mais diversos trabalhos, e dificilmente chega-se a um acordo entre essas duas correntes do cálculo das probabilidades.

Antes de *Ars Conjectandi* a matemática dos jogos de azar já existia, mas não como a probabilidade matemática que foi estabelecida por Bernoulli. Na parte II dessa obra há um tratamento sistemático da matemática de

---

[41] If the diseases multiply by the passage of time, it would be necessary to make new observations. It is certain that anyone who wanted to judge the life spans of those who lived before the flood by observations made today in London or Paris or elsewhere would stray widely from the truth.

combinações e permutações, incluindo o que veio a ser conhecido como os números de Bernoulli. Após apresentar uma série de lemas e uma rigorosa demonstração matemática Bernoulli estabeleceu na IV parte do livro o principal fundamento da sua teoria, que hoje em dia é conhecido como teorema de Bernoulli ou a lei fraca dos grandes números.

> Seja a razão entre o número de casos férteis e o número de todos os casos possíveis dado pela razão r / (r + s) ou r / t, cuja razão está limitada por (r + 1) / t e (r –1) / t. Pode-se mostrar que quanto mais experiências forem realizadas [...] provavelmente o número de observações férteis estará entre esses limites do que fora deles. (BERNOULLI, 2006, p. 337, tradução nossa) [42]

Bernoulli (2006, p.328) mostrou com esse princípio que "à medida que o número de observações aumenta, também aumenta a probabilidade de obter a verdadeira razão entre o número de casos em que algum evento pode acontecer e não acontecer." Aqui encontramos a definição de probabilidade que frequentemente é interpretada como a razão entre o número de casos favoráveis (casos férteis) e o número de todos os casos possíveis.

Após obter as bases necessárias para atribuir probabilidades a posteriori, Bernoulli pôde aplicar a matemática dos jogos de azar, por analogia, ao cálculo das probabilidades das questões civis, econômicas, demográficas e jurídicas, como nos mostra a seguinte passagem.

> Mas se isso acontecer e se, no final, a certeza moral for adquirida dessa maneira, teremos encontrado o número de casos a posteriori quase com tanta certeza como se fossem conhecidos a priori. Isso, certamente é, na prática da vida civil, [...] mais do que suficiente para dirigir nossas conjecturas em qualquer assunto contingente não menos cientificamente do que em jogos de azar. (BERNOULLI, 2006, p. 329, tradução nossa) [43]

Nesse contexto a contribuição de Leibniz, que era formado em direito, foi efetiva ao sugerir em muitas ocasiões, que alguém tratasse dos graus de probabilidades também em outras questões além dos jogos de azar. Nesse ponto podemos dizer, pelas correspondências de Bernoulli, que ele foi um continuador das ideias de Leibniz em particular na aplicação das probabilidades aos casos civis. Concomitantemente, procuramos destacar em algumas passagens da *Arte* a

[42] Let the number of fertile cases and to all the cases be in the ratio r/(r + s) or r/t, which ratio is bounded by the limits (r + l) / t and (r- 1)/ t. It is to be shown that so many experiments can be taken that it becomes [...] more likely that the number of fertile observations will fall between these bounds than outside them.

[43] But if it does happen and if in the end moral certainty is acquired in this way [...], we will have found the numbers of cases a posteriori almost as certainly as if they were known to us a priori. This, surely, in the practice of civil [...], more than suffices for directing our conjectures in any contingent matter no less scientifically than in games of chance.

importância de Leibniz na construção dessa obra, principalmente no que tange a polêmica entre estimativas de probabilidades a priori e a posteriori que culminou com a ideia fundamental de Bernoulli (2006, p.329) de que na medida em que o número de observações aumenta, aumenta a probabilidade de obter a verdadeira razão entre o número de casos em que algum evento pode acontecer e não acontecer, de modo que essa probabilidade possa, eventualmente, se aproximar da certeza. Dessa forma, Bernoulli estendeu o cálculo das probabilidades no tratamento das questões “contingentes ou livres dependendo da vontade humana ou casual dependendo da sorte ou do acaso.” (BERNOULLI, 2006, p.316):

Assim surge um esboço de um modelo estatístico bem fundamentado na premissa de que o que não pode ser verificado a priori, ou seja, por meio de estimativas iniciais baseadas no princípio da indiferença, pode ser determinado a posteriori por meio de frequências.

A Arte das Conjecturas, como sugere seu próprio nome, é apresentada por Bernoulli (2006, p. 317-318) como um método para medir probabilidades a partir de algumas regras e axiomas e, por lidar com a incerteza, essa *Arte* fornece resultados apenas prováveis. Como diria Bernoulli (2006, p.315) “A probabilidade é, de fato, um grau de certeza e difere deste último, como a parte difere do todo.”

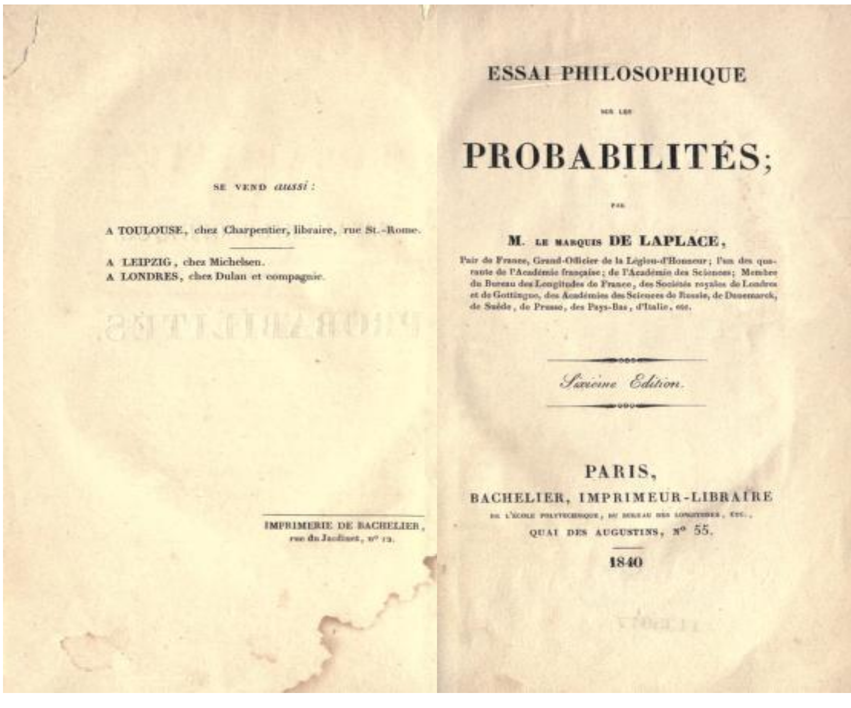
4 LAPLACE

Pierre-Simon Laplace nasceu em Beaumont-en-Auge, França, em 23 de março de 1749 e faleceu em Paris, em 5 de março de 1827. Desde cedo se interessou por vários assuntos como teologia, astronomia e principalmente matemática. Ingressou na vida acadêmica logo após ter chegado a Paris sob a recomendação de D'Alembert, que se interessou por seus trabalhos. Foi nomeado professor de matemática na Escola Militar aos vinte anos de idade e em 1773 foi aceito como membro da Academia das Ciências. Em 1785 tornou-se membro do Departamento de Geometria no novo Instituto da França. Em 1794 foi nomeado professor de análise na Escola Normal e mais tarde assumiria o cargo de presidente do *Bureau des Longitudes*. Em 1816, assumiu a presidência da comissão reorganizadora da Escola Politécnica e Membro da Academia Francesa. Recebeu de Napoleão, após o golpe de 18 de Brumário de 1779, o cargo de ministro do interior. Posteriormente recebeu de Luís XVIII o título de Marquês em 1817.

Figura 4: Retrato do Marquês de Laplace. (óleo sobre tela de 1838 do pintor por Paulin Guérin (1783-1855), Château de Versailles.
Fonte: https://fr.wikipedia.org/wiki/Pierre-Simon_de_Laplace (visitado em 20/01/2020)

Em 1796, Laplace publicou a obra *Exposição do sistema do mundo* cujo conteúdo seria incluído mais tarde no *Tratado de Mecânica Celeste* (2010, p12) que foi escrita entre os anos 1799 e 1825 em cinco volumes. Nesta obra encontram-se reunidos vários resultados surgidos no meio científico pós-newtoniano e também uma série de trabalhos originais sobre refração, pêndulos, velocidade do som e dilatação dos corpos sólidos e, principalmente, sobre a gravitação universal que era um dos assuntos de seu maior interesse.

Seus trabalhos de probabilidades concentram-se basicamente em duas obras, *A Teoria Analítica das Probabilidades* publicada pela primeira vez em 1812 e o *Ensaio Filosófico Sobre Probabilidades*, (nesse capítulo denominaremos simplesmente como *Ensaio*) (LAPLACE, 2010), cuja primeira edição é de fevereiro de 1814. Posteriormente, essa obra aparece como introdução na segunda edição da *Teoria analítica das probabilidades* (doravante *Teoria*) em novembro de 1814. O *Ensaio* foi baseado em uma palestra sobre probabilidade ministrada na Escola Normal no ano de 1795 e sofreu algumas mudanças substanciais nas três edições da *Teoria* Analítica publicadas durante a vida de Laplace.



SE VEND *aussi* :

A TOULOUSE, chez Charpentier, libraire, rue St.-Rome.

A LEIPZIG, chez Michelsen.
A LONDRES, chez Dulan et compagnie.

IMPRIMERIE DE BACHELIER,
rue du Jardinet, n° 12.

ESSAI PHILOSOPHIQUE

SUR LES

PROBABILITÉS;

PAR

M. LE MARQUIS DE LAPLACE,

Pair de France, Grand-Officier de la Légion-d'Honneur; l'un des quarante de l'Académie française; de l'Académie des Sciences; Membre du Bureau des Longitudes de France, des Sociétés royales de Londres et de Gottingue, des Académies des Sciences de Russie, de Danemarck, de Suède, de Prusse, des Pays-Bas, d'Italie, etc.

Sixième Édition.

PARIS,
BACHELIER, IMPRIMEUR-LIBRAIRE
DE L'ÉCOLE POLYTECHNIQUE, DU BUREAU DES LONGITUDES, ETC.,
QUAI DES AUGUSTINS, N° 55.

1840

Figura 5: Folha de rosto do livro Ensaio Filosófico sobre probabilidades.
Fonte: (archive.org/details/essaiphilosophiq00lapluoft)
(visitado em 10/01/2020)

Grande parte dos problemas de probabilidades abordados pelos antecessores de Laplace foram reformulados e generalizados em sua *Teoria* com o emprego do cálculo diferencial e integral, das séries e das equações diferenciais. Os princípios gerais dessa teoria, que são em número de dez, foram expostos no *Ensaio*, e em sua maioria, com exemplos de jogos de azar. Laplace não atribuiu nenhum nome aos princípios de sua teoria, mas, apesar disso, na maioria das vezes, referem-se a ele como autor do princípio da razão insuficiente. Jaynes (1957, p.622, grifo do autor), por exemplo, declarou que "o 'P*rincípio da Razão Insuficiente*' de Laplace foi uma tentativa de fornecer um critério de escolha, no qual se dizia que dois eventos devem ter probabilidades iguais se não houver nenhuma razão para se pensar ao contrário". Veremos mais adiante, nas citações de trechos da obra de Laplace, que realmente este princípio está apenas implícito em suas definições.

A ideia de uma "razão insuficiente" parece confrontar com o determinismo que é inerente ao pensamento laplaceano, como nos mostra a seguinte afirmação:

> [...] uma coisa não pode começar a ser sem que haja uma causa que a produza. Este axioma, que é conhecido pelo nome de *princípio da razão suficiente*, se aplica também as ações que julgamos indiferentes. A vontade mais livre não pode originá-las sem um motivo determinante; [...], pois, como diria Leibniz, isso seria o acaso cego dos epicuristas. (LAPLACE, 2010, p. 42, grifo nosso)

Como frisou Leibniz (1974, p.407), em várias passagens de sua obra, "há sempre uma razão para que as coisas aconteçam de um modo e não de outro". Como vimos na passagem acima, Laplace assumiu esse princípio em sua filosofia.

Posteriormente, Jaynes (1989, p. 212) também reconheceu que o princípio da razão insuficiente é uma denominação inadequada ao problema da atribuição de probabilidades iniciais. Seguindo essa afirmação de Jaynes, compreendemos que não se atribui probabilidades iguais a dois ou mais eventos por falta de uma razão, pelo contrário, a razão se encontra no reconhecimento de que esses eventos têm um mesmo número de casos favoráveis. Logo, é preciso considerar que o termo "nenhuma razão", que frequentemente aparece nos exemplos de Bernoulli, de Laplace e de Leibniz, indica que todos os eventos têm o mesmo número casos favoráveis e, consequentemente, são igualmente possíveis. Dessa forma, entendemos que o termo "nenhuma razão" implica em um problema de contagem e não de subjetividade como sugere o termo "razão insuficiente". Por exemplo, se um conjunto é formado pelos elementos {a*, a, b, b, c, c*} pode-se afirmar, com segurança, que todos os elementos deste conjunto têm o mesmo número de casos favoráveis, isto é, dois casos favoráveis. Nessas condições, não há “nenhuma razão” para atribuir mais probabilidade para a letra *c*, ou para letra *b* ou para letra *a*. Considere agora o conjunto {*a, b, b, b, c, c*}. Nessa distribuição fica claro que há um caso favorável para o elemento *a*, três casos favoráveis para o elemento *b* e dois casos favoráveis para o elemento. Como os números de casos favoráveis são diferentes, “há uma razão” para atribuir probabilidades diferentes a cada um desses casos, respectivamente 1/6; 1/2 e 1/3.

Em seu “Primeiro Princípio”[44] do *Ensaio* Laplace (1886, p.181) estabelece que “a probabilidade de um evento é a relação entre o número de casos favoráveis e o número de todos os casos possíveis, quando nada sugere que um desses casos deva ocorrer e não os outros, isso os torna para nós, igualmente possíveis.” Podemos compreender, a partir desse princípio, que para calcular a probabilidade de um evento é preciso considerar o número de todos os casos possíveis para, em seguida, identificar quantos desses casos favorecem ou não a ocorrência de um determinado evento. Matematicamente, a probabilidade corresponde a uma fração onde o numerador é o número de casos favoráveis e o denominador o número de todos os casos possíveis. Para ilustrar esse princípio vamos supor que uma urna contenha exatamente cinco fichas identificadas com as letras {a, b, c, d, e}. Nesse caso, a probabilidade de retirar qualquer uma dessas fichas é de 1/5, pois, de acordo com Laplace (1812, p.184), “se todos os casos são, a priori, igualmente possíveis, a probabilidade de cada um deles é 1/n.” Além disso, se compararmos cada uma dessas letras, sem separá-las em classes, não temos nenhuma razão para acreditar que uma delas sairá com mais frequência que as outras.

Suponhamos agora que se pretende calcular a probabilidade de se retirar da urna uma ficha identificada por uma vogal. Para calcular essa probabilidade basta considerar a própria definição de probabilidade, isto é, a razão entre o número de casos favoráveis sobre o número de todos os casos possíveis. Daí, como há dois casos favoráveis para retirar uma letra vogal, ou seja, {a, e} e cinco casos igualmente possíveis {a, b, c, d, e}, a probabilidade deste evento é, por definição, igual a 2/5.

[44] Acompanhamos a forma como Laplace intitulou os princípios do seu ensaio de probabilidades, ou seja, com as iniciais maiúsculas: Primeiro Princípio, Segundo Princípio e assim por diante.

Temos ainda uma importante propriedade que é uma consequência da definição de probabilidade de Laplace (2010, p. 46): “se o número de casos favoráveis e de todos os casos possíveis aumentam na mesma proporção, então a probabilidade permanece a mesma.”

Se acrescentarmos um caso favorável a cada um dos elementos do conjunto {a, b, c, d, e}, teremos doze casos igualmente possíveis com dois casos favoráveis para cada ficha, ou seja, {a, a, b, b, c, c, d, d, e, e}. Consequentemente a probabilidade de ocorrer qualquer um desses casos é a mesma, isto é: p(a) = p(b) = p(c) = p(d) = p(e) = 2/10 = 1/5, o que prova que a probabilidade permaneceu a mesma em relação aos elementos do conjunto {a, b, c, d, e}. Repare também que a probabilidade de se retirar uma ficha que seja uma letra vogal, não se altera no conjunto ampliado, pois temos quatro casos favoráveis para esse evento em um total de dez casos possíveis, logo p(a) = p(e) = 4/10 = 2/5.

Complementando nossa análise da definição de probabilidade e do “Primeiro Princípio” de Laplace, vamos considerar o seguinte exemplo.

> Suponhamos que se lance ao ar uma moeda larga e muito delgada cujas faces opostas são perfeitamente iguais. Procuremos a probabilidade de obter *cara* ao menos uma vez em dois lançamentos. É claro que podem ocorrer quatro casos igualmente possíveis, a saber: *cara* no primeiro e segundo lançamentos; *cara* no primeiro lançamento e *coroa* no segundo; *coroa* no primeiro e *cara* no segundo lançamento e *coroa* nos dois lançamentos. Os três primeiros casos são favoráveis ao evento cuja probabilidade é procurada, e que, consequentemente é igual a 3 / 4. (LAPLACE, 2010, pp.49-50, grifo do autor)

Esse exemplo mostra, de acordo com a definição de probabilidade e também pelo “Primeiro Princípio”, que a probabilidade de um evento corresponde à razão entre o número de casos favoráveis e o número de todos os casos possíveis, ou seja, 3 casos favoráveis em 4 possíveis. Pode-se chegar ao mesmo resultado somando as probabilidades de cada caso favorável ao evento procurado. Como temos 4 eventos igualmente possíveis, suas probabilidades são iguais a 1/4. E como temos 3 casos favoráveis para obter cara pelo menos uma vez em dois lançamentos, a probabilidade desse evento é dada pela soma de suas respectivas probabilidades, que é 1/4 + 1/4 + 1/ 4 = 3/4.

Vejamos a seguir um exemplo em que Laplace nos mostra que na medida em que o número de informações aumenta a incerteza que temos em relação à ocorrência de um evento futuro diminui e, consequentemente, maior será a probabilidade de esse evento ocorrer.

> Suponhamos, por exemplo, que se disponha de três urnas, A, B e C, de modo que uma contenha apenas bolas pretas, ao passo que as duas outras contenham apenas bolas brancas. Deve-se retirar uma bola da urna C. Pergunta-se qual a probabilidade de que essa bola seja preta. Caso se ignore qual das três urnas contém apenas bolas pretas, de modo que não haja *nenhuma razão* para se acreditar que seja a urna C e não A ou B, as três hipóteses parecerão *igualmente possíveis*; e como uma bola só pode ser extraída na primeira hipótese, a probabilidade de extraí-la é igual a 1/3. Se soubermos que a urna A contém apenas bolas brancas, a indecisão recai agora somente sobre as urnas B e C, e a probabilidade de que a bola extraída de uma urna C seja preta é de 1/2. Finalmente, essa probabilidade muda para certeza, se for garantido que as urnas A e B contenham apenas bolas brancas. (LAPLACE, 2010, p.47, grifo nosso)

Esse exemplo mostra que a equiprobabilidade (casos igualmente possíveis) está relacionada com falta de informação (nenhuma razão) e, nesse caso, a incerteza é máxima. Isso está muito próximo da afirmação de Shanonn (1948, p. 11) de que a entropia é máxima quando todas as probabilidades são iguais a 1/n e, intuitivamente, essa é a situação mais incerta.

O Sexto Princípio do *Ensaio* (2010, p.55) de Laplace tem o seguinte enunciado: “quanto mais extraordinário for um fato, mais ele precisa ser apoiado em provas sólidas, [...]”. Este princípio fornece alguns exemplos para distinguir aqueles eventos que são mais prováveis daqueles que são menos prováveis. Com isso ele procura mostrar que os eventos regulares são menos frequentes que os irregulares. Laplace exemplifica da seguinte forma: “no jogo de cara ou coroa, a combinação na qual cara ocorre vinte vezes seguidas é menos fácil que aquelas combinações em que cara e coroa aparecem intercaladas de forma irregular.”

O exemplo a seguir reflete um evento extremamente improvável ou, como se refere Laplace (2010, p.55), um evento extraordinário.

> A retirada de uma bola branca de uma urna que em 1 milhão de bolas contém apenas uma dessa cor, sendo todas as outras pretas, também parece extraordinária, pois formamos apenas duas classes de eventos relativas às duas cores.

Fica claro que a retirada de uma bola preta é extremamente mais provável do que a retirada de uma bola branca. Além disso, há uma acentuada assimetria entre esses dois eventos. Por outro lado

> a retirada do número 475.813 de uma urna que contém um milhão de números nos parece um evento ordinário, pois, comparando individualmente os números uns aos outros, sem separá-los em classes, não temos **nenhuma razão** para acreditar que um deles

> sairá preferencialmente aos outros”. (LAPLACE, 2010, p.55, grifo nosso)

Nesse exemplo está implícito a noção de uma “razão insuficiente” como um argumento para estabelecer uma igualdade de condições na retirada de um número qualquer da urna. Logo, todos os resultados têm a remota probabilidade de 1/1000000. Em outras palavras, não há nada de extraordinário na retirada do número 475.813 dessa urna, considerando que todos os resultados são, pelo princípio da indiferença, igualmente possíveis.

Nesse caso, todos os resultados são possíveis, (retirar uma bola branca ou preta) mas não são igualmente prováveis, pois a probabilidade de se extrair uma bola branca é de 1/ 1000000 enquanto a probabilidade de se retirar uma bola preta é de 999999/100000. Pode-se concluir a partir desses exemplos que quanto mais improvável for um evento, mais ele precisa ser apoiado em provas sólidas. Laplace mostra com seus exemplos que, em certas circunstâncias, o cálculo das probabilidades aplicado a esses casos também vale para outras questões que envolvem o conhecimento humano e científico. Isso fica claro em seu Terceiro Princípio quando ele mostra que a probabilidade de eventos sucessivos e independentes diminui na medida em que os eventos são repetidos. Nesse princípio são utilizados três exemplos distintos: um exemplo de jogos de dados; um relacionado à transmissão de informações pelas pessoas e outro relacionado a um fenômeno físico. Desses exemplos[45] destacamos os seguintes:

> Se os eventos são independentes uns dos outros, a probabilidade da existência de seu conjunto é o produto de suas probabilidades individuais. Assim sendo a probabilidade de obter um ás com um único dado é igual a 1/6 e aquela de obter dois ases, jogando-se simultaneamente dois dados, é 1/36. Como cada uma das faces de um dado pode se combinar às seis faces do outro, há 36 casos igualmente possíveis, dentre os quais apenas um corresponde aos dois ases. Assim, como as potências sucessivas de uma fração menor que a unidade diminui incessantemente, um evento que dependa de uma sequência muito grande de probabilidades pode tornar-se muito pouco provável. [...] Uma maneira muito interessante de comparar essa diminuição de probabilidade é observar como a nitidez dos objetos se extingue pela interposição de várias placas de vidro. Um número não muito grande de placas é suficiente para impedir a visão de objeto que uma única placa permitia perceber de maneira nítida. (LAPLACE, 2010, p.51)

Embora esses exemplos sejam de naturezas diferentes, Laplace pretende enfatizar nessa, e em outras passagens do *Ensaio*, que o cálculo das probabilidades se

[45] Consultar, https://pt.wikipedia.org/wiki/Dado_(jogo)

diversifica em várias áreas do conhecimento, sejam elas de natureza humana ou puramente científica.

Em sua concepção filosófica de probabilidade Laplace (2010, p.46) afirma.

> A probabilidade se deve em parte a nossa ignorância, em parte aos nossos conhecimentos. Sabemos que de três ou mais eventos, apenas um deve ocorrer; mas nada nos leva a crer que um deles ocorrerá preferencialmente aos outros. Nesse estado de indecisão, é impossível nos pronunciarmos com certeza sobre sua ocorrência. [...] A teoria dos acasos consiste em reduzir todos os eventos do mesmo gênero a um certo número de casos igualmente possíveis, de forma tal que estejamos igualmente indecisos sobre sua existência, em determinar o número de casos favoráveis ao evento cuja probabilidade é desejada. A relação entre esse número e aquele de todos os casos possíveis é a medida dessa probabilidade, que corresponde assim a uma fração cujo numerador é número de casos favoráveis e o denominador de todos os casos possíveis.

Essa afirmação envolve alguns obstáculos que são inerentes ao problema da atribuição de probabilidade iniciais, tais como a incerteza, a indecisão e ignorância, e a falta de um critério de escolha para distinguir, entre os vários casos possíveis, aquele que é mais provável.

Em outras passagens do *Ensaio*, Laplace procurou enfatizar a insuficiência humana para compreender os fenômenos naturais e identificar suas respectivas causas e que, na maioria das vezes, essa insuficiência nos faz remeter esses fenômenos ao acaso. Entretanto, Laplace afirmou isso fazendo uso do Princípio da Razão Suficiente de Leibniz, "*que nada acontece sem que haja uma causa que a produza*." (LAPLACE, 1825, p.42, grifo do autor) O problema é que essas causas ou essas razões nem sempre estão ao nosso alcance. Isso acontece, por exemplo, quando lançamos uma moeda sobre a mesa e o resultado é *cara* em vez de *coroa*. Na maioria das vezes desconhecemos as causas ou as razões que contribuíram para esse resultado. Poderíamos alegar uma infinidade de circunstâncias tais como, a força atribuída à moeda no momento do lançamento, o atrito da moeda com a superfície da mesa, uma possível assimetria entre as faces da moeda e daí por diante. Seria necessária uma inspeção minuciosa para que pudéssemos atingir a última razão dentro de uma série de acontecimentos. Mas isso não implica em dizer que não houve uma causa ou uma razão que tenha contribuído para esse resultado.

Há nas estimativas de probabilidades iniciais uma tendência em querer projetar o futuro no presente, mas essa projeção torna-se mais difícil quando essas estimativas estão relacionadas aos fenômenos naturais que frequentemente envolvem tantas circunstâncias e um grande número de causas que é quase inconcebível fazer alguma previsão dentro dos limites dos nossos conhecimentos. De acordo com Laplace[46] isso só seria possível para

> uma inteligência que, em um dado instante, conhecesse todas as forças que animam a natureza com seus respectivos seres e, além disso, fosse suficientemente ampla para submeter todos esses dados à análise, nada lhe seria incerto e o futuro bem como o passado estariam presentes em seus olhos [...] O espírito humano só oferece um frágil esboço dessa inteligência. (LAPLACE, 2010, pp. 42, 43)

Em virtude de que desejamos mostrar a importância que Leibniz teve, historicamente, nestas concepções, o que faremos mais detalhadamente no próximo capítulo, podemos complementar essa passagem do *Ensaio* com o seguinte pensamento de Leibniz (1974, p.119-121):

> Pode-se dizer em consequência das pequenas percepções, que o presente é grande e que o futuro está carregado do passado, [...]. Todavia, compete à suprema razão, à qual nada escapa, compreender distintamente todo o infinito e enxergar todas as consequências. Tudo o que podemos, com respeito às grandezas infinitas é conhecê-las confusamente, e saber ao menos que elas existem.

Independente da natureza do evento, seja ele de origem natural ou em um simples lançamento de uma moeda, haverá sempre três momentos distintos, um antes um durante e outro depois.

Poderíamos ainda destacar outros pontos de convergência entre as filosofias desses dois pensadores que acentua ainda mais essas aproximações e que de certa forma mostra que, em ambos os casos, há pensamento determinista em suas filosofias, principalmente quando Laplace assume, em sua filosofia das probabilidades, o Princípio da Razão Suficiente de Leibniz.

Finalmente, podemos afirmar que Laplace elaborou uma teoria das probabilidades em toda sua amplitude, isto é, estabeleceu princípios filosóficos e sistematicamente os demonstrou em sua análise matemática. Por outro lado, sabe-se que Leibniz não construiu uma teoria das probabilidades nos mesmos padrões estabelecidos por Laplace. No entanto, encontra-se em seus manuscritos e também em algumas passagens de sua obra filosófica, uma considerável inserção do seu pensamento no cálculo das probabilidades.

---

[46] Posteriormente esse pensamento de Laplace ficou denominado, metaforicamente, como “demônio de Laplace”.

5 LEIBNIZ

Leibniz nasceu em 1646 na cidade de Leipzig e ingressou na Universidade Leipzig em 1661 onde obteve os graus de bacharel e mestre em filosofia, respectivamnete em 1663 e 1664. Na Universidade de Altdorf obteve no ano de 1666 o título de doutor em direito com a tese “De casibus perplexis in jure[47]”. No início de sua formação acadêmica, ele vislumbrou a possibilidade de construir uma linguagem que pudesse ser entendida por todos e aplicada aos mais diversos ramos do conhecimento. A construção dessa linguagem foi iniciada em sua “Dissertação sobre a Arte Combinatória”[48] onde encontram-se relacionados os conteúdos de filosofia, de lógica e de aritmética. Ele acreditava que, através dessa linguagem, que posteriormente chamou de “Característica Universal”, todo o pensamento humano poderia ser construido a partir de combinações e de permutações de algarismos e letras que são elementos fundamentais ou os termos simples de uma proposição.

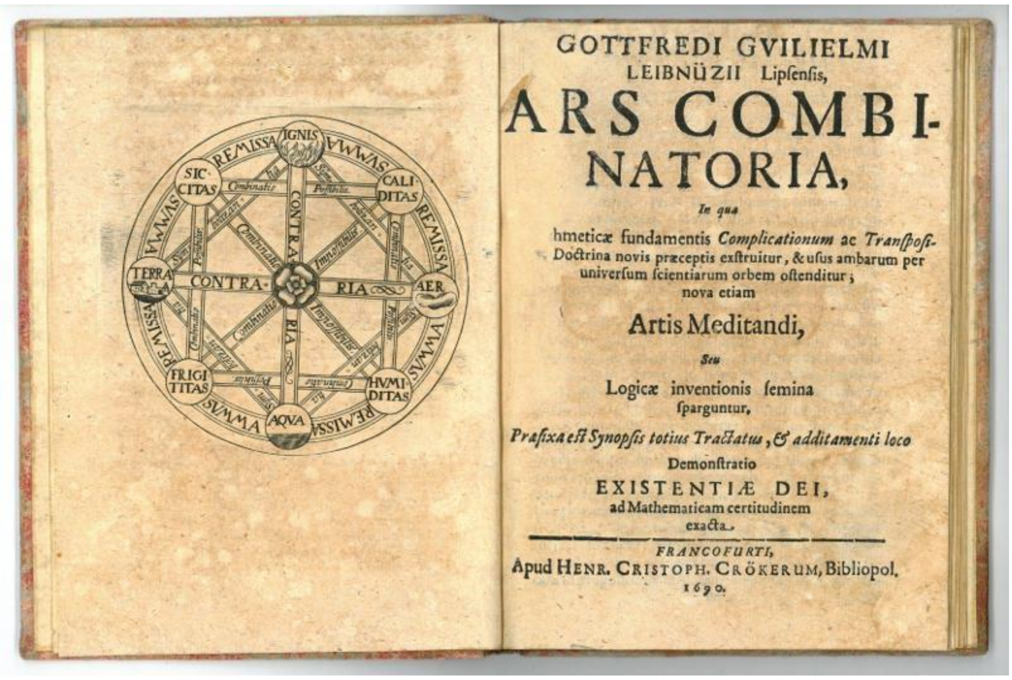


Figura 6: frontispício e a folha de rosto da Arte Combinatória de Leibniz
Fonte: https://pictures.abebooks.com/INLIBRIS/22423011450.jpg (visitado em 12/12/2019)

Destacamos essa obra não somente por sua importância para seus trabalhos posteriores de séries, combinações e probabilidades mas também porque a ideia de uma linguagem universal, ou seja, uma linguagem acessível a todos, refletiu em

[47] Traduzido para o Inglês por E. J. Aiton, como “On difficult cases in law”. (Sobre casos difíceis da Lei)
[48] Foi escrita originalmente em latim com o título “Dissertatio de Art Combinatoria” no ano de 1666.

outras invenções, como foi o caso do Cálculo Infinitesimal[49] que, com o emprego das notações dy e dx para representar os "elementos infinitesimais ou diferenciais" (EDWARDS, 1937), possibilitou resolver vários problemas que envolviam, por exemplo, o cálculo de áreas (quadraturas) e de tangentes. Desse modo, Leibniz destacou-se no meio científico do século XVIII também pela generalização que apresentou em seus métodos matemáticos e por suas notações que possibilitaram generalizar vários problemas resolvendo-os analiticamente, quer dizer, pelo uso de fórmulas ao invés de figuras. Como o próprio Leibniz costumava afirmar "Minhas soluções são sempre universais" (LEIBNIZ,1962, pp.182-190). Esse pensamento que permeou grande parte de suas invenções foi um dos fatores que contribuíram para o seu ingresso definitivo na comunidade científica daquela época, mais precisamente como um membro da Royal Society of London onde havia um certo ceticismo quanto aos seus conhecimentos iniciais de matemática, principalmente por parte dos cientistas ingleses, que tinham em Newton a sua maior referência científica.

Em seu processo de construção o Cálculo Infinitesimal de Leibniz teve sua origem em seus trabalhos iniciais de sobre as séries numéricas[50] e também de combinações. Como se sabe, esses dois ramos da matemática encontram-se também no cálculo das probabilidades. A exemplo disso temos a distribuição binomial que nos fornece a probabilidade de um evento aleatório ocorrer ou não ocorrer em uma série de tentativas. Esse estreito enlace entre esses ramos da matemática, fica mais evidente em alguns manuscritos de Leibniz sobre jogos onde as combinações servem para definir, por exemplo, o número de todos os resultados possíveis quando dois ou mais dados são lançados. Leibniz frequentemente enfatizava a necessidade de incluir a soma das séries no cálculo das probabilidades e foi o que fez Bernoulli em sua teoria das probabilidades. (ver anexo B)

Os estudos das séries já tinha sido fundamental para o Cálculo Infinetsimal como pode-se compreender na seguinte passagem da carta que Leibniz enviou a Abbé Conti no ano de 1716.

> Foi reunindo minhas antigas observações sobre as diferenças das séries dos números com as minhas novas meditações de geometria

[49] A construção do cálculo infinitesimal de Leibniz começou em 1673 e concretizou-se no ano de 1675 e foi publicado em 1684. Sugerimos como consulta para esses fatos históricos a obra de Margareth Baron, "The Origins of the Infinitesimal Calculus". Oxford, Pergamon, 1969.

[50] Para maiores informações consultar "O Pensamento Inicial de Leibniz sobre as Séries e o Método das Diferenças". Antunes, Marcelo Mattos. 2005, UFRJ.

> que eu encontrei, aproximadamente em 1676, um novo cálculo, que chamei de Cálculo das Diferenças, cuja aplicação à geometria produz coisas maravilhosas. (LEIBNIZ, 1986, p.98-104, tradução nossa)[51]

Na introdução da "Arte Combinatória", Leibniz refere-se[52] a algumas obras e também a alguns cientistas que, possivelmente, lhe serviram de fonte de consulta em seus primeiros trabalhos de séries, combinações e probabilidades. Entre essas referências, encontra-se o nome do dinamarquês Erasmus Bartholin e do holandês Frans van Schootem que foi aluno de Descartes e professor de Huygens e que escreveu, no ano de 1657, o livro *Exercitarum Mathematicanum*, que inclui questões de aritmética e de geometria. Também cita os *Elementos* de Euclides, que foi traduzido por Isaac Barrow. Sobre as combinações, consultou o livro intitulado *Recreações Matemáticas* de Daniel Schwenter e o texto elementar de astronomia, que também envolve combinações, *La Sphère* de Jean Sacrobosco. Finalmente, menciona no texto o livro *Arithmétique Pratique* do matemático italiano Jérôme Cardan e o livro *Nova Scientia* de Nicolo Tartaglia que, como se sabe, contribuiram significativamente para o desenvolvimento do cálculo das probabilidades. Posteriormente, esses conhecimentos iniciais de combinações (consultar apêndice A), permutações e séries infinitas foram aplicados em seus manuscritos de jogos e também nas estimativas de pensões vitalícias e de expectativas demográficas. Dentre esses manuscritos temos: "Cálculo das Partes" (Sur le calcul des partis-1676); "Estimativas da incerteza" (*De incerti Aestimatio*-1678); "Jogo do cinco e nove" (*Du jeu Qinquenove*-1678); "Expectativa de vida" (*Lebenserwartung*) que também envolve o cálculo integral; "Rendas vitalícias" (*De Reditibus ad vitam*-1680) "Meditações jurídico-matemática sobre os interesses intermediários" (*Meditatio juridico mathematica de interusurio simplice*-1683) com o envolvimento de números figurados e de combinações associadas às séries infinitas. Desse modo, as ideias de Leibniz sobre probabilidades encontram-se fragmentadas em seus manuscritos[53] e também dispersas em suas obras filosóficas e correspondências.

---

[51] C'est en rassemblant mes anciennes observations sur les différences dans les séries de nombres avec mes nouvelles méditations géométriques que j'ai trouvé, vers 1676, un nouveau calcul, que j'ai appelé Le Calcul des différences, dont l'application à la géométrie produit des choses merveilleuses.

[52] As referências destacadas por Leibniz encontram-se nas seguintes páginas da Arte Combinatória: 113,115,117, 118, 122 e 144.

[53] Sobre os jogos de dados Leibniz escreveu dois manuscritos com os títulos "Nombre de faces dans les lancers de dés" (1676) e "Du jeu de Quinquenove" (1678). Sobre jogo de cartas "Du jeu de la Bassette" (1678) e mais três manuscritos com os títulos "Le jeu du Solitaire" (1678); "Divinations arithmétiques", "Cálculo das Partes" e "Note sur certains jeux" (1710).

Entre os manuscritos mencionados acima cabe destacar *De incerti Aestimatio* (Estimativas da incerteza), pois nele Leibniz define o conceito de probabilidade que foi traduzido e sintetizado pelo filósofo e matemático Louis Couturat (1969, p. 245) da seguinte maneira:

> As probabilidades estão para a certeza como as partes estão para o todo ou como a fração própria está para a unidade. Com efeito, a probabilidade de um evento sendo definida pela relação entre o número de casos favoráveis pelo número de todos os casos possíveis, só pode ser uma fração própria, e quando esta é igual à unidade, a probabilidade torna-se uma certeza. Para isso, é preciso que todos os casos sejam igualmente possíveis.[54]

Observamos na passagem acima que a definição de probabilidade, dada pela razão entre o número e de casos favoráveis e o de todo os casos possíveis, envolve, implicitamente, a ideia de uma equiprobabilidade ou de uma igualdade entre todos os casos possíveis que é a condição necessária para determinar o número de casos favoráveis. Além disso, essa passagem nos faz relembrar Bernoulli (2006, p.315), quando este define que a “probabilidade, de fato, é um grau de certeza e difere da própria certeza como a parte difere de todo”. E também Laplace (1886, p.181) quando estabelece que “a probabilidade de um evento é a relação entre o número de casos favoráveis e o número de todos os casos possíveis, quando nada sugere que um desses casos deva ocorrer e não os outros, isso os torna para nós, igualmente possíveis.” Mais uma vez vemos nessas definições uma correspondência entre as ideias de desses três pensadores em particular a ideia de uma distribuição de probabilidades iguais (casos igualmente possíveis) que se encontra subentendida em suas definições de probabilidade, como vimos nos dois capítulos anteriores. Adiante veremos como a ideia de uma equiprobabilidade se justifica em Leibniz pelo princípio da indiferença.

Dependendo do contexto em que se encontra inserido, o conceito de probabilidade em Leibniz pode receber diferentes denominações como “aparência”, “esperança” ou “receio” [55]. A “aparência” que geralmente foi a denominação mais utilizada em seus manuscritos de jogos de azar, é composta de “esperança” e “receio” que designam, respectivamente, a “probabilidade de ganhar” e a

---

[54] Les probabilités sont à la certitude comme les parties sont au tout ou comme les fractions propes à l’unité. Et en effet, la probabilité d’um événement étant définie le rapport du nombre des cas favorables au nombre de tous cas possibles, ne peut être qu’une fraction propre, et quand celle-ci est égale à l’unité, la probabilité devient une certitude. Pour cela, il faut que tous les cas soient également possibles (tradução nossa).

[55] O termo “aparência” encontra-se no manuscrito *L’estime de L’incertain* e “esperança e receio” no manuscrito *Du jeu de quinquenove.* (Leibniz, 1995)

"probabilidade de perder". Esses termos, que também foram aplicados em suas abordagens aos problemas de expectativa de vida e de estimativas de rendas vitalícias, refletem, sobretudo, a incerteza que é inerente ao próprio conceito de probabilidade, como nos sugere o seguinte pensamento de Leibniz (1999, p. 688, tradução nossa)

> Participamos de certas atividades e também de alguns jogos, mesmo sabendo que não há segurança, pois há uma ciência que nos governa dentro das próprias incertezas e nos permite descobrir de que lado há mais aparência. Mas é surpreendente que essa ciência seja quase desconhecida e que os lógicos ainda não tenham examinado os graus de probabilidades para fornecer uma estimativa tão precisa quanto os números. Entretanto, essa estimativa não pode e não deve ser usada para chegar a uma certeza, o que é impossível, mas para agir de maneira mais razoável possível sobre os fatos ou conhecimentos dados a nós. [...] Existe, portanto, uma ciência sobre os assuntos mais incertos, que torna conhecido, demonstrativamente, os graus de aparência e da incerteza [56]

Compreendemos que esse pensamento forneceu, desde sua época, um conteúdo bastante importante às afirmações probabilísticas atuais na ciência. Também fica claro nesse pensamento a relação entre probabilidade e incerteza o que nos faz lembrar a afirmação de Jaynes (1957) de que a única maneira de quantificar a incerteza é através de uma distribuição discreta de probabilidades e que uma ampla distribuição representa mais incerteza do que uma precisão acentuada. Apesar de se tratar de situações completamente distintas, a incerteza e a imprevisibilidade, que envolvem as estimativas de probabilidades iniciais, não foi um problema exclusivo dos jogos de azar e nem dos cálculos de expectativa de vida pois, conforme já mencionamos anteriormente, esse problema também foi um tema central no MaxEnt, talvez até com maior frequência e por isso cabe discuti-lo aqui.

Como sugeriu Leibniz (1986, p.457) em algumas passagens de seus manuscritos,[57] fazer estimativas diante da incerteza é querer projetar o futuro no presente e, para esses casos, deve-se sempre tomar o valor médio entre várias

[56] C'est de ces emplois comme du jeu, ou il faut se resoudre et prendre party lors même qu'il n'y a nulle asseurance, il y a une science qui nous gouverne dans les incertitudes mêmespour de´couvrir de quel costé la plus grande apparence se trouve. Mais il est étonnant qu'elle est presque inconnue et que les Logiciens n'ont pas encor examiné les degrés de probabilité ou de vraisemblance qu'il y a dans les conjectures ou preuves qui ont pourtant leur estimation aussi asseurée que les nombres; cette estimation nous peut et doit servir non pas pour venir à une certitude, ce qui est impossible mais pour agir le plus raisonnablement qu'il se peut sur les faits ou connoissances qui nous sont données.[...] Il y a donc une science sur les matieres les plus incertaines, qui fait connoistre demonstrativement les degrés de l'apparence et de l'incertitude.

[57] "O capital atual é sempre equivalente a todas as rendas futuras ao infinito, considerando todos os tempos como presentes" (Leibniz, 1680, p. 377)

suposições igualmente possíveis. Segundo o historiador Parmentier (1995, p. 341) a originalidade dessa regra baseia-se na ideia de uma equiprobabilidade e, eventualmente, sobre o princípio da razão insuficiente. Parmentier (1995, p.36-37) ainda acrescenta que "a legitimidade de uma estimativa inicial se fundamenta sobre um princípio de indiferença." Em várias passagens de suas obras Jaynes também relaciona, alternadamente, o princípio da razão insuficiente e o princípio da indiferença ao problema das estimativas iniciais de probabilidades uniformes. Assim, tanto Jaynes como Parmentier nos sugerem que essas duas denominações são equivalentes. Todavia, seguiremos a sugestão de Keynes de que a denominação princípio da indiferença é aquela que melhor se ajusta ao problema da atribuição de probabilidades iguais. No decorrer desta seção procuraremos justificar essa nossa opção através do pensamento leibniziano, mais precisamente em seu conceito de indiferença.

Até aqui destacamos alguns fatos históricos que precederam os trabalhos de probabilidades de Leibniz e a maneira como esses trabalhos foram estabelecidos. Também mostramos que há em seu pensamento uma relação entre probabilidade e incerteza que é uma das questões epistemológicas presentes no MaxEnt.

A seguir destacaremos alguns exemplos dos manuscritos de Leibniz para que possamos identificar e discutir os argumentos e os princípios leibniziano que envolvem o problema da atribuição de probabilidades iguais.

Conforme vimos nos exemplos de jogos de Bernoulli, só é possível atribuir probabilidades iguais se soubermos, *a* priori, que há um argumento para isso ser feito, como nos mostra o exemplo abaixo.

> Quando lançamos um dado, todos os resultados têm *chances iguais* para ocorrer; por causa da semelhança de suas faces e do seu peso uniforme, sendo assim *não há razão* para que uma das faces seja mais propensa a cair do que outra. (BERNOULLI, 1713, p.326, grifo nosso, tradução nossa).[58]

Esse suposto equilíbrio, isto é, a uniformidade do dado, também aparece como uma condição necessária para justificar uma distribuição probabilidades iguais no seguinte exemplo de Leibniz (1986, p.457, grifo nosso, tradução nossa).

> A aparência é o grau de probabilidade, tomemos o seguinte exemplo, se um dado tem suas seis faces equilibradas, cada uma delas terá a mesma aparência de ocorrer, ou seja: não há *nenhuma*

[58] So, for example, the numbers of cases in dice are known. Moreover these all have equal tendencies to occur; because of the similarity of the faces and the uniform weight of the die, there is no reason why one of the faces should be more prone to fall than another.

> *razão* pela qual se pode dizer, que a face 1 ocorrerá mais do que a face 2, ou 3, ou 4, ou 5, ou 6. [59]

Vimos nos dois exemplos acima que a informação de que o dado é equilibrado em relação à sua forma geométrica é, sob o ponto de vista objetivo, a razão ou o argumento necessário para atribuir probabilidades iguais a todas as faces do dado. Por outro lado, a afirmação de que não há nenhuma razão para atribuir mais probabilidade a uma face em vez de outra qualquer reflete, sob o ponto de vista subjetivo[60], os limites de nossas compreensões e a nossa incapacidade de perceber que, no momento do lançamento do dado, há uma infinidade de circunstâncias, como os micromovimentos e as pequenas disposições que nos são imperceptíveis, que são os componentes que determinarão qual das faces ficará voltada para cima quando o dado ficar em repouso.

Em seu Princípio da Razão Suficiente Leibniz (2017, p. 161) afirma que "jamais algo acontece sem que haja uma causa ou uma razão determinante [...] embora, muito frequentemente, essas razões não nos sejam conhecidas o suficiente, nós não deixamos de pressentir que elas existem." Ao mesmo tempo em que o Princípio da Razão Suficiente aponta para a necessidade de uma razão em tudo o que acontece, ele também revela a nossa insuficiência para compreender toda a série acontecimentos que precedem um determinado evento. É essa incapacidade de compreender esse encadeamento de causas que nos leva a afirmar que não há nenhuma razão (apesar de existirem) para que um evento ocorra em vez de outro qualquer. Em suma, dizer que não há razão para um evento acontecer não significa dizer, pelo princípio da razão suficiente, que tais razões não existem e, devido a esse determinismo, Leibniz (1995, p.446, tradução nossa) criticou o seguinte pensamento de Jean Le Clerc:[61]

> O senhor Le Clerc diz que a combinação dos bilhetes de loteria é efeito de um movimento que damos a eles para misturá-los; e que esse movimento é o efeito de uma inteligência livre, que agita mais ou menos as bolas ou os bilhetes, mais pelo acaso do que pela razão, e que não há nenhuma razão que explique o movimento

---

[59] l'Apparence n'est autre chose que le degre´ de laprobabilité; par exemple un dé, dont on se sert pour jouer, ayant ses six faces bien egales, l'apparence est egale pour chacune de ces faces, c'est à dire, il n'y a point de raison pour la quelle on puisse dire, qu'il tombera plus tost un 1, qu'un 2, ou 3, ou 4, ou 5, ou 6.

[60] Segundo Jaynes (1957, p.622) "A escola do pensamento "subjetivo" considera as probabilidades como expressões da ignorância humana; a probabilidade de um evento é meramente uma expressão formal de nossa expectativa de que um evento acorreu ou ocorrerá, com base em alguma informação que esteja disponível. Para o subjetivista, o propósito da teoria da probabilidade é nos ajudar a tirar conclusões plausíveis nos casos em que não há informação."

[61] Esse trecho foi extraído do manuscrito de Leibniz intitulado *Sur les Loteries*, publicado em 1696, onde Leibniz fez uma crítica a obra de Jean Le clerc, *Réflexion sur ce que l'on appelle bonheur et malheur en matière de loteries et sur bon usage qu'on en peut faire,* (Parmentier, 1995, p. 443).

> dado a esses bilhetes e nem porque agitar a caixa dez vezes, por exemplo, e não onze.[62]

Entretanto, Leibniz refuta esse pensamento de Le Clerc na seguinte passagem de sua crítica.

> É verdade que ele não é obrigado a dar razão ao número de movimentos**,** mas isso não impede que essas razões ou essas causas existam. [...] Mas podemos dizer ainda que isso acontece também nas causas não mecânicas, ou seja, aquelas que envolvem uma escolha, pois essa escolha sempre tem suas razões perceptíveis ou imperceptíveis que vêm do nosso entendimento ou dos nossos órgãos ou dos objetos. (LEIBNIZ, 1995, pp. 446-447, tradução nossa) [63]

O pensamento exposto acima pode ser compreendido como um desdobramento do Princípio da Razão Suficiente de Leibniz que fundamenta sua filosofia e também grande parte de sua obra científica, por exemplo, em seus trabalhos de dinâmica. Conforme destacamos anteriormente no capítulo 4, Laplace reafirma a importância desse princípio na seguinte passagem de *Ensaio*.

> Os eventos atuais têm com os precedentes uma ligação fundamental no princípio evidente de que nada pode acontecer sem que haja uma causa que a produza. Esse axioma, conhecido como *princípio da razão suficiente* se estende até mesmo às ações que julgamos indiferentes. A vontade mais livre não pode originá-las sem um motivo determinante. (LAPACE, 2010, p.42, grifo do autor)

Essa indiferença a qual se referiu Laplace [64] na citação acima significa, de acordo com Leibniz (2017, p. 346; 347) que

> somos livres para escolher, mas jamais escolhemos quando não vemos nenhuma diferença entre as partes. [...] e, embora nem sempre veja a razão que me faz escolher entre duas partes que parecem iguais, sempre existirá algum detalhe, embora imperceptível, que nos determina nessa escolha.

Para ilustrar esse princípio, Leibniz utilizou o clássico exemplo da balança[65] de Arquimedes, através do qual compreendemos que há implicações análogas

---

[62] M. Le Clerc dit que la combinaison des billets de loterie est l'effet d'un mouvement que nous leur donnons pour les mélanger; et que ce mouvement est l'effet d'une intelligence libre, qui secoue plus ou moins les boules ou les billets, plus par hasard que pour cause, et qu'il n'y a aucune raison d'expliquer le mouvement donné à ces billets ou pourquoi secouer la boîte dix fois , par exemple, pas onze.

[63] Il est vrai qu'il n'est pas obligé de rendrer raison au nombre de secousses, mais cela n'empêche point qu'il y en ait des raisons ou ces causes. [...] Mais on peut dire qu'il n'en est pas autrement dans des causes non machinales où il entre du choix parce que ce choix a toujours ses raisons perceptibles ou imperceptibles qui viennent de de l'entendement ou des organes ou de l'objet.

[64] Não podemos dizer com precisão até que ponto Leibniz influenciou Laplace em sua filosofia das probabilidades mas, como já destacamos anteriormente, esse pensamento e também em outras passagens do seu *Ensaio,* encontramos uma acentuada aproximação entre as ideias desses pensadores.

[65] O exemplo da balança foi frequentemente utilizado por Leibniz para ilustrar seu princípio da razão suficiente, tanto em sua filosofia quanto em sua física. Apesar de se tratar de um caso particular, esse exemplo serve, segundo Leibniz, para "demonstrar a divindade e o resto da metafísica, e também os princípios físicos independentes da matemática, isto é, os princípios dinâmicos, ou da força." (Leibniz, *Correspondências com Clarke*, p. 408, 1974).

àquelas que estão inseridas nos exemplos dos jogos de dados e também das moedas:

> se houver uma balança em que tudo seja igual nos dois lados, e se também suspendermos pesos iguais nas duas extremidades desta balança, tudo permanecerá em repouso. Isso ocorre porque *não há razão* para que um lado desça mais do que o outro. (LEIBNIZ, 1974, p.407)

Seguindo essa linha de raciocínio também podemos pensar em um dado ou em uma moeda que, se tiverem suas faces bem equilibradas, não haverá razão para que ela se incline mais para um lado do que para outro, desde que não haja nenhuma interferência externa. Essa interferência externa pode ser, por exemplo, a força atribuída à moeda no momento em que ela é arremessada e atinge o solo. É na composição desses fatores externos e outros que não percebemos que faz a moeda inclinar para um lado e não para o outro. Sendo assim, o mais razoável é atribuir a mesma probabilidade para as duas faces da moeda, a saber, 1/2. Nesse caso, podemos perguntar o que leva alguém escolher, por exemplo, cara em vez de coroa já que se sabe, a priori, que a moeda é, por hipótese, equilibrada? Aparentemente não há uma razão para essa escolha mas, de acordo com Leibniz,

> há uma infinidade de grandes e de pequenos movimentos internos e externos que ocorrem conosco, dos quais o mais frequentemente é não se aperceber; quando saímos de um quarto existem certas razões que nos determinam a colocar um ou outro pé na frente, sem que se reflita sobre isso. (LEIBNIZ, 2017, p. 162)

O problema de uma suposta simetria perfeita nas formas geométricas, em particular, nos dados ou nas moedas, foi detalhadamente discutido por Jaynes (2003, p. 314) pois, além das questões epistemológicas desse problema que são inerentes ao cálculo das probabilidades, também envolve outras questões de natureza estritamente técnicas, como ressalta Jaynes na seguinte passagem.

> as situações em que temos conhecimento positivo de simetria são bastante especiais entre todos aqueles enfrentados pelo cientista. Como podemos realizar um raciocínio indutivo consistente em situações em que não percebemos nenhum elemento claro de simetria? Este é um problema aberto porque há uma infinidade de circunstâncias especiais que podem surgir. Como veremos, o princípio da máxima entropia fornece uma ferramenta útil e versátil para muitos desses problemas. Mas, para começar a entender isso, vamos voltar ao início e considerar o lançamento da moeda mais uma vez. (JAYNES, 2003, p. 335, tradução nossa)[66]

---

[66] in situations where we do not perceive any clear element of symmetry? This is an openended problem because there is no end to the variety of different special circumstances that might arise. As we shall see, the principle of maximum entropy gives a useful and versatile tool for many such problems. But in order to give a start toward understanding this, let's go way back to the beginning and consider the tossing of the coin still another time.

A passagem acima nos mostra que até nas questões mais elementares, como é o caso dos jogos de dados e de moedas, qualquer tentativa de atribuir probabilidade por meio de frequências nos envolve em algumas dificuldades lógicas se levarmos em conta as causas mecânicas do experimento que ocasionaram, por exemplo, uma moeda cair com uma face voltada para cima em vez da outra. Se soubermos, através de medições físicas, que o dado é perfeitamente simétrico e aceitamos as leis da mecânica como corretas, deduzimos que a variação dos resultados, correspondentes às faces do dado, é decorrente de uma não uniformidade no método de lançamento. Se não conseguimos, por algum motivo particular, repetir um experimento sob as mesmas condições iniciais, por exemplo, imprimir a mesma força de arremesso no momento do lançamento de uma moeda, espera-se, por consequência, que tenhamos diferentes resultados nessas condições. Desse modo as conclusões que podemos tirar do experimento aleatório depende de sabermos ou não que o dado é simétrico.

Por outro lado, o reconhecimento de um elemento de simetria entre as faces do dado nos conduz, pelo Princípio de Indiferença, a uma atribuição de probabilidades uniformes. Como ressaltou Jaynes (1957, p. 622), essa simetria "restringe-se a poucas situações". Leibniz é ainda mais enfático quando afirma, em seu princípio dos indiscerníveis que, "*há sempre um pormenor que faz a diferença e, em virtude das variações insensíveis, duas coisas individuais não podem ser completamente semelhantes.*" (LEIBNIZ, 1974, p. 120)

Apesar de reconhecer que tudo acontece por uma a razão e que "o acaso é apenas algo aparente, como a sorte (*fortune*); pois, é a ignorância das causas que o faz." (LEIBNIZ, 2017, pp. 346-347), Leibniz também reconhece que há uma sequência de razões que são imperceptíveis para nós, pois há sempre um pormenor que nos escapa. Esse detalhe ou esse pormenor que desconhecemos no momento do lançamento de um dado, nos leva a um suposto estado de indiferença ou de hesitação para escolhermos um ou outro resultado e, nesse estado de indiferença, podemos suspender nossas ações ou nossas escolhas.[67] Como diria Leibniz (2017. p.346), "jamais escolhemos quando somos absolutamente indiferentes e, apesar de sermos livres para escolher, há sempre uma razão que nos conduz em nossas escolhas". Se alguém nos informa que um dado tem pesos iguais em todas as suas faces, isto é: que ele é equilibrado, a única escolha que

[67] *Discurso de Metafísica* (Leibniz, 1974, p. 103).

nos resta é atribuir probabilidades iguais a todas elas, caso contrário teríamos uma escolha arbitrária. Entretanto, essa simetria ou essa indiferença de equilíbrio é uma situação excepcional, limitando-se a um domínio restrito das probabilidades dadas *a priori* como ocorre, por exemplo, no contexto geométrico dos jogos de dados onde há a possibilidade de uma simetria e de um equilíbrio supostamente perfeitos.

Vejamos agora com um exemplo de Leibniz onde a probabilidade de um evento depende do número de casos que lhe são favoráveis.

> Mas se jogarmos dois dados ao mesmo tempo e combinarmos o número de pontos dos dois dados a fim de obter uma soma, haverá mais probabilidade para que a soma seja igual a sete pontos do que doze. Além disso, a probabilidade de fazer 7 pontos é o triplo da probabilidade para fazer 12 pontos. Pois, para fazer 12 pontos há apenas uma maneira, isto é 6 e 6; enquanto que pode-se fazer 7 pontos de três maneiras *igualmente prováveis*, ou seja: 6 e 1, 5 e 2, e 4 e 3. (LEIBNIZ, 1986, p.457, tradução nossa) [68]

Nesse exemplo, Leibniz considerou apenas três resultados favoráveis para a soma sete, que são os pares (6,1); (2, 5) e (3, 4) e não considerou as possíveis permutações desses resultados que são (1, 6); (5, 2) e (4, 3). Para a soma doze temos apenas um caso favorável para esse evento que é (6, 6) e, nesse caso, a permutação não gera um novo resultado. Se essas possíveis permutações não forem consideradas em cada par ordenado, teremos, de acordo com o diagrama de Leibniz, apenas 21 casos possíveis

(1; 1), (1; 2), (1; 3), (1; 4), (1; 5), (1; 6)

(2; 2), (2; 3), (2; 4), (2; 5), (2; 6)

(3; 3), (3; 4), (3; 5), (3; 6)

(4; 4), (4; 5), (4; 6)

(5; 5), (5; 6)

(6; 6)

No diagrama acima, Leibniz considerou apenas as combinações com repetição de seis elementos tomados dois a dois, ou seja:

---

[68] Mais si l'on jette deux dés à la fois, et qu'on assemble le nombre des points des deux dés pour en faire une somme, il y aura plus d'apparence, qu'on fera sept points que douze. Et même l'apparence de faire 7 points est triple de l'apparence qu'ily a d'en faire 12. Car il n'y a qu'une seule maniere de faire 12 points, en jettant 6 et 6; mais on peut faire 7 points en trois manieres egalement faisables, par 6 et 1, par 5 et 2, et par 4 et 3.

$(CR)_{6;\ 2}$ =, $C_{6+2-\ 1;\ 2}$ = $C_{7;\ 2}$ = 21 resultados possíveis, com quinze combinações formadas por elementos distintos, por exemplo, 1 e 4, 3 e 6, 2 e 5,... e seis combinações com elementos repetidos, que são: 1 e 1 , 2 e 2, 3 e 3, 4 e 4, 5 e 5 e 6 e 6. Logo, temos um total de 21 resultados possíveis todos com a mesma probabilidade de 1 / 21. De acordo com o exemplo acima, a probabilidade de que tenhamos a soma 7 no lançamento simultâneo de dois dados será igual à soma dos casos que lhe são favoráveis pelo total de casos possíveis, ou seja, P(7) = 3 / 21 ou, proporcionalmente falando, 1 caso favorável em cada 7 possíveis. Como temos um caso favorável para somar 12 pontos, a probabilidade desse evento é p(12) = 1/ 21. Mas se considerarmos as possíveis permutações entre os pares ordenados acima, excetuando os resultados duplos (1; 1), (2; 2), (3; 3), (4; 4), (5; 5), (6; 6), teremos um total de 36 resultados possíveis. Dessa forma, teremos seis casos favoráveis para que soma seja 7, isto é: (1; 6), (6; 1), (2; 5), (5; 2), (3; 4), (4; 3) e, consequentemente, a probabilidade desse evento será igual a 6 / 36 ou, proporcionalmente falando 1 caso favorável em cada 6 possíveis. Vimos assim que sua probabilidade aumentou contrariando assim a afirmação de Leibniz de que a probabilidade para fazer 7 pontos é o triplo da probabilidade de se fazer 12 pontos.

Naturalmente, na medida em que aumentamos o número de dados a complexidade e a imprevisibilidade dos resultados aumentam proporcionalmente e, consequentemente, a probabilidade de cada resultado diminui. Com um dado temos um total de seis casos possíveis com probabilidades iguais a um 1/6. Com dois dados temos 36 resultados possíveis com probabilidades iguais a 1/36 e para três dados temos 216 casos possíveis com probabilidades iguais a 1/216 e assim sucessivamente. De uma forma geral o número total de casos possíveis corresponde ao número de arranjos com repetição, ou seja, $A(r)_{n,\ p} = n^p$, onde n é o número de elementos ou de resultados possíveis e p indica quantos elementos participaram em cada subconjunto. Por exemplo, com um dado temos n = 6 e p= 1, logo $A(r)_{6,1} = 6^1$= 6 casos possíveis. Para dois dados temos n= 6 e p = 2, logo $A(r)_{6;\ 2} = 6^2$ = 36 casos possíveis. Para três dados temos n = 6 e p= 3, logo temos $A(r)_{6;\ 2} = 6^3$ = 216 casos possíveis.

Vamos analisar agora um exemplo semelhante ao anterior onde encontramos implícito o princípio da indiferença como um argumento para atribuir probabilidades iguais.

> Suponha que eu jogue dois dados e que esses dados sejam bem feitos [...]. Desse modo, existem apenas duas maneiras para fazer

> 5 pontos: uma é 1 e 4 e a outra 2 e 3 e, por outro lado, há três maneiras para fazer 8 pontos, a saber, 2 e 6; 3 e 5 e 4 e 4. Cada uma dessas maneiras tem a mesma probabilidade, pois *não há nenhuma razão* que nos permita dizer que há mais probabilidade de ocorrer como resultado 1 e 4 do que 3 e 5. (LEIBNIZ, 1995, p.195, grifo nosso, tradução nossa) [69]

Para calcular a probabilidade de cada resultado, basta contar o número de casos favoráveis e dividir pelo número de casos possíveis. Assim, como há duas "combinações ou maneiras" (LEIBNIZ, 1995, p.202) para somar cinco pontos em vinte e um casos possíveis, a probabilidade desse evento é 2/21. Do mesmo modo temos três casos favoráveis para somar oito pontos, logo a probabilidade desse caso é 3/21. Este exemplo faz parte do jogo de dados chamado Cinco e Nove. Para jogá-lo é preciso lançar dois dados. Se na primeira jogada sair dois números repetidos (1, 1), (2, 2), (3, 3), (4, 4), (5, 5) ou (6, 6) o jogador ganha. Mas também pode ganhar se somar 3 pontos ou 11 pontos que acontece, isto é, (1, 2) e (5, 6). Como já falamos anteriormente, Leibniz considerou um total de 21 resultados possíveis, logo há oito graus de probabilidade para um jogador vencer na primeira jogada, ou seja, 8/21. Por outro lado, o jogador perde o jogo se soma dos pontos for igual a 5 ou 9 (daí o nome do jogo), que corresponde aos pares (1, 4), (2, 3), (3, 6), (4, 5), assim há quatro graus de probabilidade para um jogador perder, ou seja, 4/21. Os resultados restantes são indiferentes, pois o jogador não perde e também não ganha, a saber, (1, 3), (1, 5), (1, 6), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (4, 6), então há nove graus de indiferença para a primeira jogada, cuja probabilidade é dada por 9/21 ou 3/7. Como nesse caso o jogador não ganha e nem perde, o jogo continua e o jogador só vencerá, a partir da segunda jogada, se ele fizer o mesmo número de pontos da primeira jogada. Essas são as regras básicas desse jogo.[70]

Vimos no exemplo de Leibniz (1995, p.195) que a simetria ou um suposto equilíbrio entre as faces dos dados é a condição necessária para atribuir a mesma probabilidade a cada um dos possíveis resultados. Em contrapartida, essa simetria ou esse suposto equilíbrio torna qualquer resultado imprevisível já que todos têm a mesma chance de ocorrer. De acordo com Leibniz (2017, p. 165), "não há nada que torne um evento necessário quando existe a possibilidade de ocorrer outro

---

[69] Supposons que je lance deux dés, et que les dés soient bien faits, s'il y a des tours, il est évident qu'il n'y a que deux façons de faire 5 points: l'un est 1 et 4 et l'autre 2 et 3, d'autre part, il y a trois façons de faire 8 points, un sabre 2 et 6; 3 et 5 et 4 et 4. Chacun de ces moyens a la même probabilité, car il n'y a pas de raison de dire ce qui est là, mais la probabilité de se produire en résultat 1 et 4 que 3 et 5.

[70] Para uma compreensão mais detalhada dessas regras e da inserção das séries numéricas e das combinações nos jogos de dados sugerimos os seguintes manuscritos de Leibniz (1995): *Nombre des faces dans les lancers de dés* (1676) e *Du jeu de Quinquenove* (1678).

evento em seu lugar.” Podemos dizer com alguma segurança que esse princípio se aplica a diversos jogos de azar cujos resultados não são previsíveis. Em outros termos, tudo que é “acidental”,[71] “mutável”[72] ou que “não corresponde a nenhuma regra lógica” [73] é contingente. Porém, esses termos não retratam toda a complexidade que envolve o conceito de contingência que Leibniz aborda em grande parte de suas obras filosóficas. Não discutiremos nesse estudo os vários desdobramentos desse conceito. Nos interessa com esse estudo revelar algumas de suas implicações no contexto das probabilidades e procurar mostrar que o conceito de indiferença está fundamentado no “princípio da contingência[74] que, segundo Leibniz (1974, p.87), “garante a existência e a possibilidade da escolha entre diversas coisas igualmente possíveis”.

Leibniz deixa claro em várias passagens de sua obra que indiferença significa o mesmo que contingência, como nos mostra o seguinte trecho de sua *Teodicéia*: “são contingentes ou indiferentes todos os eventos que independem de qualquer necessidade lógica e não tem nada em si que possa conferir um princípio de certeza” (LEIBNIZ, 2017, p. 160). Em outros termos, a contingência garante a pluralidade dos eventos ou da existência das coisas, ou seja, no que é ou parece ser o melhor entre várias coisas igualmente possíveis. Além disso, a escolha de um ou outro evento qualquer não acarreta nenhuma contradição, caso contrário, o evento seria necessário.[75] Assim, quando lançamos uma moeda supostamente simétrica, não é necessário sair cara em vez de coroa considerando que os dois resultados são igualmente possíveis. Portanto, “não há nada que torne um evento necessário quando se pode pensar que qualquer outra coisa pudesse acontecer em seu lugar”. (LEIBNIZ, 2017, p.165).

Mas é preciso levar em conta que indiferença significa o mesmo que contingência desde que a entenda como uma livre escolha entre diversos casos igualmente possíveis. Logo, não se trata de uma indiferença vulgar ou absolutamente indeterminada pela qual se escolhe sem nenhum motivo ou nenhuma razão, pois “jamais escolhemos quando somos absolutamente

[71] Leibniz (1974, p.307).
[72] Leibniz (1974, p.393).
[73] Esses termos encontram-se dispersos nas obras de Leibniz, por exemplo em *Origens Primeira das Coisas* (1974, p.393) e também na *Teodicéia* (2017, 338).
[74] Leibniz (1970, p.87)
[75] “há duas espécies de verdades: as de *Razão* e as de *Fato*. As verdades de Razão são necessárias, e o seu oposto, impossível; as de Fato, são contingentes, e o seu oposto, possível. Quando uma verdade e necessária pode encontrares-lhe a razão por meio da Analise, decompondo-a em ideias e verdades mais simples, até alcançar as primitivas.” (Leibniz, 1974, p.66, grifo do autor)

indiferentes. Tal escolha seria uma espécie de puro acaso, sem razão determinante, tanto uma que apareça quanto uma escondida." (LEIBNIZ, 2017, p. 346). Contudo, podemos perguntar: o que leva alguém escolher entre dois eventos que, aparentemente, encontram-se em igualdade de condições? Como um exemplo mais específico, o que leva alguém escolher uma das faces de uma moeda quando se sabe, a priori, que a moeda é equilibrada ou simétrica em relação às suas faces? Será que em toda escolha há um motivo ou uma razão que a determine? De acordo com Leibniz (1974, p.380), "parece impraticável que uma pessoa se incline mais para o lado em que vê menor probabilidade". Mas o que acontece quando nada indica que uma das possibilidades é mais provável do que a outra, como proceder nesse caso? A resposta para essas perguntas, dependerá do conceito de indiferença em Leibniz que ele distingue em dois níveis distintos.

No primeiro nível, encontra-se a indiferença de equilíbrio, que Leibniz rejeita, que nos leva a indecisão ou a uma escolha feita ao acaso sem nenhuma razão determinante.[76]

No segundo nível, há uma indiferença, que ele admite, pela qual é possível escolher livremente e diversamente entre vários casos possíveis. Mas apesar dessa liberdade, toda escolha se faz por uma determinada razão e isso exclui a possível "quimera de uma indiferença de equilíbrio"[77], ou seja, "uma escolha que seria espécie de puro acaso, sem razão determinante."[78]

Vemos então que a concepção de indiferença em Leibniz se encontra polarizada. Diferentemente do que se pensa no senso comum, o conceito de indiferença em Leibniz não se resume a um descaso a um desinteresse ou, em termos leibniziano, a uma indiferença de equilíbrio. Essa oposição estabelecida por Leibniz entre indiferença de equilíbrio e indiferença enquanto uma livre escolha se justifica no princípio da razão suficiente que, em uma das suas versões tem o seguinte enunciado:

> há sempre algo de verdadeiro que determina a nossa vontade para *escolha* do que nos parece melhor, mas desde que isso não seja necessário. Pois, absolutamente falando, a vontade possui uma *indiferença* desde que ela se oponha à necessidade, e tem o poder de *escolher* diversamente ou ainda de suspender inteiramente sua ação, pois ambas as coisas são e continuam sendo *igualmente possíveis*. (Leibniz, 1974, p.103, grifo nosso)

76 Leibniz (2017, p.346)
77 Idem.
78 Leibniz (2017, p.347)

Nota-se nessa passagem que o conceito de indiferença e seus vínculos com o problema da contingência tem implicações filosóficas de extrema complexidade, como é o caso do livre arbítrio, por exemplo. Entretanto, uma discussão mais aprofundada desse assunto não está dentro dos nossos propósitos.

Além dessas questões de cunho estritamente filosófico que acabamos de mencionar, o problema da indiferença também tem implicações em problemas das ciências físicas e matemáticas com abordagens probabilísticas, como pode ser visto nos textos de probabilidades de Jaynes, por exemplo. Nesse contexto, nos chamou a atenção a seguinte pergunta de Jaynes (2003, p.573): “a indiferença é baseada no conhecimento ou na ignorância?”

Foi exatamente em Leibniz (1989, p. 488) onde encontramos duas possíveis respostas para essa pergunta, ou seja: “a indiferença surge da ignorância, e quanto mais sábio se é, mais se está determinado a fazer o que é mais perfeito.”

Em outra passagem Leibniz (1989, p. 337) afirma que:

> Uma indiferença perfeita é uma suposição enganosa ou incompleta. Ver-se-á que tiro conclusões surpreendentes desse princípio declarado, mas isso ocorre apenas porque não costumamos levar suficientemente longe nosso conhecimento mais claro.

A indiferença a qual Leibniz (2017, p.156) se refere nessa passagem é a indiferença de equilíbrio ou uma indeterminação, como se estivéssemos igualmente inclinados para dois lados diferentes. Como já ressaltamos anteriormente, essa indiferença deve ser rejeitada, pois esse equilíbrio perfeito ou essa simetria são apenas “ficções que já mais ocorre no universo, de modo que tudo seja igual e semelhante em ambos os lados”. [79] Mas devemos lembrar mais uma vez que há uma indiferença que nos permites escolher livremente entre vários casos possíveis. Entretanto, essa escolha é sempre conduzida por uma razão determinante que nos tira de uma condição de equilíbrio e nos faz inclinar mais para um lado do que para o outro.

Consideramos que essas duas concepções antagônicas da indiferença e, sobretudo as passagens que destacamos anteriormente, contribuíram para atender a nossa expectativa de que há, explicitamente, um Princípio de Indiferença em Leibniz, como esperávamos mostrar nessa tese. Em suma, encontramos em elementos suficientes para que possamos reafirmar que os fundamentos epistemológicos do Princípio da Máxima Entropia, que antes estavam relacionados

[79] Leibniz (2017, p. 163)

somente a Bernoulli e a Laplace, também podem ser discutidos em Leibniz através de sua obra filosófica, sobretudo, o princípio da indiferença para o qual encontramos uma sólida fundamentação filosófica.

Encerramos esse capítulo destacando a importância de Leibniz, segundo o parecer de Keynes (1921, p. 313, tradução nossa).[80]

> Não seria justo deixar de falar do grande Leibniz, que, mais sábio em correspondências e projetos fragmentários do que em discursos completos, deixou-nos indicações suficientes de que suas reflexões particulares sobre esses assuntos estavam muito adiantadas em relação aos seus contemporâneos. Ele distinguiu três graus de convicção entre opiniões, certeza lógica [...] e certeza física que nada mais é que a probabilidade lógica.

---

[80] It would not be just here to pass by entirely the name of the great Leibniz, who, wiser in correspondence and fragmentary projects than in completed discourses, has left to us sufficient indications that his private reflections on this subject were much in advance of his contemporaries'. He distinguished three degrees of conviction amongst opinions, logical certainty [...] and physical certainty which is only logical Probability.

CONCLUSÕES

Inicialmente procuramos mostrar a importância do MaxEnt para vários ramos da física e também para diversas áreas do conhecimento científico. Além de ser bastante simples na sua formulação matemática, esse princípio oferece alguns recursos muito interessantes por ser um método puramente estatístico, ou seja: é um método que independe de suposições físicas e tem a característica de fornecer conclusões confiáveis a partir de poucas informações.

Procuramos compreender por que o MaxEnt foi considerado uma generalização do princípio da razão insuficiente. Vimos que esse princípio serviu apenas como um ponto de partida para Jaynes elaborar um método estatístico para atribuir probabilidades.

Apesar de vários autores relacionarem o princípio da razão insuficiente ao nome de Laplace, constatamos que essa denominação não corresponde a nenhum dos dez princípios do seu *Ensaio filosófico de probabilidades*. Também procuramos identificar nessa obra os princípios e os conceitos que levaram a ideia de uma distribuição de probabilidades uniformes. Concluímos que essa ideia se encontra na própria definição de probabilidade de Laplace.

Seguindo as afirmações de Jaynes (1978, p. 212), investigamos na *Arte das Conjecturas* os princípios estabelecidos por Jacobi Bernoulli em suas estimativas de probabilidades. Vimos que essas estimativas estão fundamentadas nessa obra por meios de combinações e permutações e suas diversas aplicações nos jogos de azar. Também constatamos, através de suas correspondências[81] e bibliografias, a importante contribuição de Leibniz para a teoria das probabilidades de Bernoulli.

Procuramos compreender como o problema da atribuição de probabilidades iniciais se encontrava nos manuscritos de probabilidades de Leibniz. Em seguida, discutimos o importante conceito de indiferença em Leibniz que, por sua vez, possibilitou analisar o problema da simetria como um critério de escolha para atribuir probabilidades. Mostramos que o conceito da indiferença em Leibniz está vinculado aos princípios da contingência, da razão suficiente e dos indiscerníveis. Sendo assim, destacamos a importância que Leibniz teve nas ideias que levaram ao princípio da indiferença e assim contribuir para esclarecer algumas questões do MaxEnt, como o problema das informações *a priori* que frequentemente são

[81] Sobre as correspondências trocadas entre Leibniz e Bernoulli, sugerimos como fonte de consulta a obra Edith Sylla, "The Emergence of Mathematical Probability from the Perspective of the Leibniz-Jacob Bernoulli Correspondence," *Perspectives in the Sciences* 6 (1998): 41-76.

abordadas por diversos autores. Essas questões estiveram muito presentes na obra de Leibniz, cuja contribuição nos parece bastante evidente, apesar de não encontramos citações de sua obra relacionadas aos fundamentos da MaxEnt.

Também destacamos alguns exemplos e conceitos da obra de Leibniz que são muito semelhantes aqueles abordados por Bernoulli e Laplace. No caso de Bernoulli, as cartas trocadas por ambos, revelam claramente a influência de Leibniz sobre ele, conforme já ressaltamos anteriormente. Vimos também em algumas passagens dos *Ensaios* de Laplace que há uma forte aproximação com o pensamento filosófico de Leibniz. Contudo, sabemos que Laplace construiu sua Teoria das Probabilidades com o uso da análise matemática, enquanto Leibniz direcionou seu conhecimento inicial das probabilidades para problemas particulares dos jogos, das questões do direito e das estimativas das rendas vitalícias. Mas foi possível encontrar nos manuscritos de Leibniz de probabilidades, alguns conceitos que foram fundamentais tanto para a teoria Laplace quanto para a teoria de Bernoulli, por exemplo, a inserção das séries no cálculo das probabilidades.

Através dessa tese vimos que outras frentes de pesquisas também podem ser exploradas a partir do problema da atribuição de probabilidades e de seus respectivos princípios. Seguindo a sugestão de Jaynes (1957), pode-se elaborar um estudo mais aprofundado sobre o problema da especificação de probabilidades confrontando a escola de pensamento "objetivo" [82] [83] com a escola do pensamento subjetivo. No primeiro caso considera-se a probabilidade de um evento como resultado de experimento aleatório sempre capaz de ser medido por frequências relativas. Por outro lado, a escola do pensamento "subjetivo"[84],[85] considera a probabilidade como expressão da ignorância humana e a probabilidade é meramente uma expressão formal da expectativa que se tem sobre a ocorrência de um evento com base em alguma informação que esteja disponível.

Os manuscritos de Leibniz englobam uma série de questões que podem revelar as origens do cálculo estatístico que o sucederam. Nesse contexto, sugerimos uma inspeção nos manuscritos sobre pensões e rendas vitalícias e de expectativa de vida. Nesses dois casos, as séries infinitas servem como elemento fundamental para projetar o valor futuro de uma renda ou a idade média de uma população. Em suma, sugerimos que uma análise sobre os manuscritos e os

[82] S. H. Cramer, Métodos de Estatística (Princeton University Press, Princeton, 1946).
[83] W. Feller, Introdução à Teoria da Probabilidade e suas aplicações (John Wiley and Sons, Inc., Nova Iorque, 1950).
[84] J. M. Keynes, 2 Treatise on Probabitity (MacMillan Company, London, 1921).
[85] H. Jeffreys, Teoria da Probabilidade (Oxford University Press,1939).

diversos assuntos que ali se encontram poderá evidenciar alguns fatos relevantes para a história do desenvolvimento inicial do cálculo das probabilidades e da estatística abrindo diversas perspectivas para novas pesquisas.

REFERÊNCIAS

ABBOTT, B. P. et al. 2016. **Phys. Rev**. Lett. 116 (6): 061102.

ABLES, J. G. 1974, **Astron. Astrophys**. *Suppl*, 15, p. 383.

AITON, E. J. **Leibniz – A Biography**. Oxford: Pergamon, 1969.

ANDREI, A.; COELHO, B.; GUEDES, L.; LYRA, A. 2019, **Monthly Notices of the Royal Astronomical Society**, v. 488, p.183-190, **The Principle of Maximum Entropy and the Luminosity Function of Quasars**.

ANTUNES, M. M. **O Pensamento Inicial de Leibniz sobre as Séries e o Método das Diferenças.** Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em História das Ciências e das Técnicas e Epistemologia. 2005.

ANTUNES, M. M.; LYRA, A. **O Princípio da Máxima Entropia e o Problema da Razão Insuficient***e*. XI Scientiarum do HCTE/UFRJ; 2018.

ANTUNES. M. M.; LYRA, A.; KOEHLER, C. **The Genesis of the "Principle of Insufficient Reason" in Leibnizian Thought and its Implications in the Principle of Maximum Entropy** – 25th International Congress of History of Science and Technology, Rio de Janeiro, 23 to 29 July 2017.

ARIEH, B. N. **A Farewell to Entropy**: *Statistical Thermodynamics Based on Information*. *Published by* World. Copyright © 2008 by World Scientific Publishing Co. Pte. Ltd. USA

ATICHA, A.; PREUSS, R. 2004, **Phys. Rev**. E, v.70, 046127, **Maximum entropy and Bayesian data analysis**: Entropic prior distributions.

BAÑADOS, E. et al. 2018, **Nature**,553, 473B.

BARON, M. **Origins of the Infinitesimal Calculus**. Oxford, Pergamon, 1969.

BEKENSTEIN, J. D. 1974, **Phys. Rev**. D, v9, N12, p .3292-3300, **Generalized second law of thermodynamics in black-hole physics**.

______. 1975, **Phys. Rev**. D, v.12, N.10 **Statistical blackhole thermodynamics**, p.3077.

BERNOULLI, J. **The Art of Conjecturing together with Letter to a Friend on sets in Court Tennis**-Translated with an introduction and notes by Edith Dudley Sylla - The Johns Hopkins Urtiversity Press. Printed in the United States of America. Published. 2006.

BREWER, B. J, 2008, Lett. to **Nature**, https :// letterstonature:wordpress:com. 2008 Where do I stand on maximum entropy.

BOREL, E. **Valeur Pratique et Philosophie des P***robabilités*. Gauthier-Villars, Imprimeur-èditeur, de L'École Polytechnique, 55, Quai des Grandes-Augustins, 55. 1939.

BOS, H. J. M.; BARON, M. **Curso de História da Matemática: O Cálculo no Século XVIII**. Unidade 4. Brasília ed. Universidade de Brasília, 1985**.**

CARNAP, R. **Logical Foundations of Probability**. Printed by the University of Chicago Press; Chicago, *Illinois*, U.S.A. Second Impression 1963.

CHILD, J. M. **The Early Mathematical Manuscripts of Leibniz**. Chicago, Open Court,1920.

COUTURAT, L. **La Logique de Leibniz**, d'après des documents inédits, Paris, Félix Alcan, 1901, reprint Hildeshein, 1969.

COSTABEL, P. **Leibniz et les Séries Numériques**. In: **Leibniz a Paris** *(1672 1678).* Symposion de G.W. Leibniz–(Hannover) et du Centre National de la Recherche Scientifique (Paris) a Chantilly (France) du 14 au 18 November 1976. Tome I, Les Sciences, ed. Frans S. V. Gmsh. Wiesbaden, 1978.

CRAMER, H. **Mathematical Methods of Statistics**. Priceton University Press, Princeton, 1946.

DELEUZE, Gilles, **A Dobra: Leibniz e o Barroco**.Trad. Luiz B.L. Orlandi. Campinas, 2000.

DEWAR, R. C. 2005, **Maximum entropy production and the fluctuation theorem**. J. Phys. A: Math. Gen., v38, L371-L381

DEWAR, R. C.; MARITAN A. 2014, **A Theoretical Basis or Maximum Entropy Production**. p.49, in Beyond the Second Law Entropy Production and Non-equilibrium Systems, Edt Dewar.

EDWARDS, A.W. F. **Pascal's arithmetical triangle.** London: C. Griffin, New York: Oxford University Press, 1987. 174 p.

EDWARDS, C. H. **The Historical Development of the Calculus**, Springer-Verlay, N. York, 1937.

FÉRON, C.; HJORTH, J. 2008, **Phys. Rev**. E, v77, 022106. **Simulated dark-matter halos as a test of nonextensive statistical mechanics**.

GULL, S. F.; DANIELL, G. J. 1978, **Nature**, 272, 686-690

HACKING, I. **The Emergence of Probability** - A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference- Second Edition Cambridge University Press. 2006.

HALD, A. **A History of Probability and Statistics and Their Applications before 1750.** New York: Wiley, 1990.

HOFMANN, J. **Leibniz in Paris** – *1672-1676*. London, 1901.

JAYNES, E. T. **Information Theory and Statistical Mechanics I**, Phys. Rev., Vol. 106, p. 620 and Vol. 108, 1957.

______. **Where do we Stand on Maximum Entropy**. The Maximum Entropy Formalism, A Conference Held at the Massachusetts Institute of Technologyon, may 24. Edited by Raphael Do Levine and Myron Tribus. 1978.

______.**The Minimum Entropy Production Principle**. 1980. Ann. Rev. Phys. Chem., 31, 579-601.

______.**Probability Theory the Logic of Science**. Edited by G. Larry Bretthorst. Cambridge University Press, New York. 2003.

JEFFREYS, H. **Theory of Probability**, Clarendon Press. Oxford University Press 1961.

______.**Scientific Inference.** Cambridge University Press. Cambridge. 1957

KEYNES, J. M. **A Treatise on Probability**. Maccmillan and Co. pp.41-64. 1921.

KONDEPUDI, D.; PRIGOGINE, I. 1998, **Modern Thermodynamics**, p. 392, Wiley, New York.

LAPLACE, P.S. **Théorie Analytique des Probabilités**. Paris, Gauthier-Villars et editors, Troisième Édition,1820. Ouvres Complètes.

______. **Ensaio Filosófico sobre as Probabilidades**. Tradução, introdução e notas Pedro Leite de Santana- Rio de Janeira: Contraponto: Ed. PUC-Rio, 2010.

LEIBNIZ., G. W. **Dissertation Sur L'art Combinatoire**. In: *Ouvre Mathematique*, v.1, Jean Pyroux, Bordeaux Janvier, fascicule I, pp. 111-169, Paris1986.

______. **Dissertation Arithmetique Sur Les Complexions**.In: *Ouvre Mathematique*, v.1, Jean Pyroux, Bordeaux Janvier, fascicule I, pp. 111-169, Paris1986.

______. **Historia et Origo Calculi Differentialis**. In: *Mathematische Scriften,* C. Gerhardt, Halle, 1850-1853, reprint George Olms, v.5, pp. 392-410, 1962.

______. **Novos Ensaios Sobre o Entendimento Humano**. Victor Civita, São Paulo, 1974.

______. **Recherches Générales sur L'Analyse des Notions et des Vérites**. 24 theses métaphysiques et autres textes logiques et métaphysiques. Introductions et notes par Jean-Baptiste Rauzy, PUF, Paris, 1998.

______. **La Monadologie**. von, edit. C. J. Gerhardt, vol. VI, Die philosophischen Schriften,1961.

______. **Die philosophischen Scriften**, Berlin, 1875 -1890, réédition Olms, 1978.

______. **Ensaios de Teodiceia- sobre a bondade de Deus, a liberdade do homem e a origem do mal**. São Paulo-SP: Ed. Estação Liberdade Ltda, 2017.

______. **Discours de Métaphysique**. Librairie Philosophique J. Vrin 6, Place of the Sorbonne, Ve, 1994.

______. **Philosophical Papers and Letters**. Dordrecht: Kluwer Academic Publishers,1989.

______. **Philosophische Schriften,** Vierter Band - Akademie Verlag, 1999.

______. **Essay de Quelques Raisonnemens Nouveaux sur la Vie Humaine et sur le Nombre des Hommes**. Sämtliche Schriften und Briefe, Akademie-Verlag Berlin, pp. 1986.

______. **Lettre de Leibniz and Galloys**. In: Mathematische Scriften, C. Gerhardt, Halle, 1850-1853, reprint George Olms, v.I/II, pp. 182- 190, 1962.

______. **Da Origem Primeira das Coisas**. In: Pensadores, Victor Civita, São Paulo, 1974.

______. **Correspondência com Clarke**. In: Pensadores, Victor Civita, São Paulo, 1974.

______. **L'estime des apparences. 21 manuscrits de Leibniz sur les probalités, la théorie des jeux, l'espérance de vie**. Texte établi, introduit par J. Echeverría, traduit, annoté et postfacé par M. Parmentier, VRIN, Paris, 1995.

______. **Carta de Leibniz para Abbé Conti**, **9 / 4 / 1716**. In: **Ouvres concernant Le calcul infinitésimal**. v.1, Jean Pyroux, Bordeaux Janvier, pp. 98-104, 1986.

HEIDELBERGER, M. **Origins of the logical theory of probability**: Von Kries, Wittgenstein, Waismann, International Studies in the Philosophy of Science, 15:2, 177-188. 2001.

MARTYUSHEV, L. M.; SELEZNEV, V. D. **Phys. Rev**. v.426, 2006, p.1-45, **Maximum entropy production principle in physics chemistry and biology**.

MOHAMMAD-DJAFARI, A.; DEMOMENT, G. 1988. in Skilling J. ed., **Maximum Entropy and Bayesian Method***s.* Springer, Cambridge, p.195.

MOSTELLER, F.; ROUKER, R. E. K.; THOMAS, G.B. **Probability with Statistical Applications**, Addison Wesley Publ. Co., Reading-Mass. 1965.

PASCAL, B. **Traité de Triangle arithmétique**, Paris, Ouvres II, pp. 445 -503. 1665.

______. PASCAL, B.; **Great Books of the Western Worl**d - The Provincial Letters Pensees Scientific Treatises, V. 33, 1952.

PONTZEN, A.; GOVERNATO, F. **Monthly Notices of the Royal Astronomical Society**, v.430, 2013, pages 121-133.

PRIGOGINE, I., **Introduction to Thermodynamics of Irreversible Process.** 3rd Ed. John Wiley & Sons, New York Completar, 1967.

______. **Time, Structure, and Fluctuations,** Science, vol. 201, No. 4358, 1978, pp. 777-785.

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

RICHARDS, G. T. et al. (2006) **Astrophys***.* J, v.131, 2766.

SAVAGE, L. J*.* **Foundations of Statistics**, Dover, New York, 1954.

SERRES, M. **Le Système de Leibniz**. Presses Universitaires de France, 1ª ed., Paris,1968.

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

SHIMONY, A. **Phys. Rev. Lett. Synthese**, v. 63, p-35-53 (1985).

SKILLING, J; BRYAN, R. K. 1984, **Mon. Not. R. Astron. Soc**., 211, 111-124.

SYLLA, E. **The Emergence of Mathematical Probility from the Perspective of the Leibniz- Jacob Bernoulli Correspondence**- Perspectives in the Sciences 6; 41-76. 1998.

TIKOCHINSKY, Y.; TISHBY, N. Z.; LEVINE, R. D. (1984) **Phys. Rev**. Lett. v. 52, p.1357, **Consistent Inference of Probabilities for Reproducible Experiments***.*

UFFINK, J. **Can the maximum entropy principle be explained as a Consistency requirement***?* Stud. Hist. Phil. Mod. Phys. V.26, n.3, p. 223-261. 1995

WALPOLE, R. E. et al. **Probabilidade e Estatística**. 8. ed. Tradução Luciane F. Pauleti Vianna – São Paulo: Pearson Prentice Hall, 2009.

ZIEGLER, H., **An Introduction to Thermomechanics**. North-Holland Publishing Company, Amsterdam, 1987. p. 229.

ZUNCKEL, C.; TROTTA, R. 2007, **Monthly Notices of the Royal Astronomical Society**, v.380, p. 865.

APÊNDICES

APÊNDICE A

O dispositivo elaborado por Leibniz para calcular todas as combinações que podem ser formadas a partir de um determinado número termos é a tabela das complexões (Tabela Aleph), onde vemos intervir, pela primeira vez, o pensamento sobre *séries* na *Arte Combinatória*. Posteriormente esses conhecimentos iniciais de séries e combinações fizeram parte dos manuscritos de Leibniz sobre o cálculo das probabilidades.

Tabela ℵ

	0	1	1	1	1	1	1	1	1	1	1	1	1	1	
	1	0	1	2	3	4	5	6	7	8	9	10	11	12	← **Números**
e	2	0	0	1	3	6	10	15	21	28	36	45	55	66	
x	3	0	0	0	1	4	10	20	35	56	84	120	165	220	
p	4	0	0	0	0	1	5	15	35	70	126	210	330	49	
o	5	0	0	0	0	0	1	6	21	56	126	252	462	792	
e	6	0	0	0	0	0	0	1	7	28	84	210	462	924	
n	7	0	0	0	0	0	0	0	1	8	36	120	330	792	
t	8	0	0	0	0	0	0	0	0	1	9	45	165	49	
e	9	0	0	0	0	0	0	0	0	0	1	10	55	220	
s	10	0	0	0	0	0	0	0	0	0	0	1	11	66	
	11	0	0	0	0	0	0	0	0	0	0	0	1	12	
	12	0	0	0	0	0	0	0	0	0	0	0	0	1	
	*	0	1	3	7	15	31	63	127	255	511	1023	2047	4095	
	†	1	2	4	8	16	32	64	128	256	512	1024	2048	4096	

* Complexions simplement (soma das complexões para um expoente dado);
† acrescentando a unidade, coincide com os termos de uma progressão geométrica dupla.

Na tabela **א**, as complexões encontram-se distribuídas em treze colunas (da esquerda para direita, excetuando-se a coluna dos expoentes) e nas treze primeiras linhas de cima para baixo de 0 a 12. Cada coluna corresponde às complexões de um número da segunda linha da tabela, a saber: 0, 1, 2,..., 10, 11, 12. Assim a primeira coluna corresponde às complexões do *número* 0, que são todas nulas, pois para os casos em que o *expoente* é superior ao *número* dado, o valor da complexão é sempre zero. Por exemplo, com um conjunto de três elementos só podemos formar subconjuntos de um, dois e três elementos. A segunda coluna é formada pelas complexões do *número* 1, a terceira coluna pelas complexões do *número* 2, e assim sucessivamente. Por outro lado, a primeira linha corresponde ao *expoente* 0, a segunda linha, ao *expoente* 1, a terceira linha, ao *expoente* 2, e assim por diante. Isto significa dizer que a primeira linha é formada por combinações do tipo $C_{n,0;}$ a segunda linha por combinações do tipo $C_{n,1;}$ a terceira linha por combinações do tipo $C_{n,2}$ e assim sucessivamente. Sendo assim, para encontrar as complexões de um determinado *número*, basta relacioná-lo com os expoentes menores ou iguais a ele. Por exemplo: o *número* 4 da quinta coluna com o *expoente* 0 da primeira linha corresponde à unidade; o *número* 4 com o *expoente* 1 formam 4 *uniões* ou quatro combinações de um elemento; o *número* 4 com o *expoente* 2 formam 6 *com2nações* ou seis combinações de dois elementos; o *número* 4 com o *expoente* 3 formam 4 *com3nações* ou quatro combinações de três elementos e o *número* 4 com o *expoente* 4 formam uma *com4nação* ou uma combinação de quatro elementos. Logo, a coluna correspondente ao *número* 4, (quinta coluna) é formada pelas complexões 1, 4, 6, 4, 1. Utilizando a notação atual temos: $C_{4,\,0} = 1$, $C_{4,\,1} = 4$, $C_{4,\,2} = 6$, $C_{4,\,3} = 4$ e $C_{4,\,4} = 1$. Em outros termos, podemos dizer que as complexões correspondentes a um determinado *número* estão localizadas nas interseções de sua coluna com as respectivas linhas dos expoentes da tabela. Sendo assim, a complexão localizada na linha de expoente 0 e na coluna do número 4 é igual a 1; a complexão localizada na linha do expoente 1 e na coluna do número 4 é igual a 4; a complexão localizada na linha do expoente 2 e na coluna do número 4 é igual a 6 e assim por diante. Do mesmo modo, as demais complexões estão distribuídas pelos *números* e pelos respectivos *expoentes*.

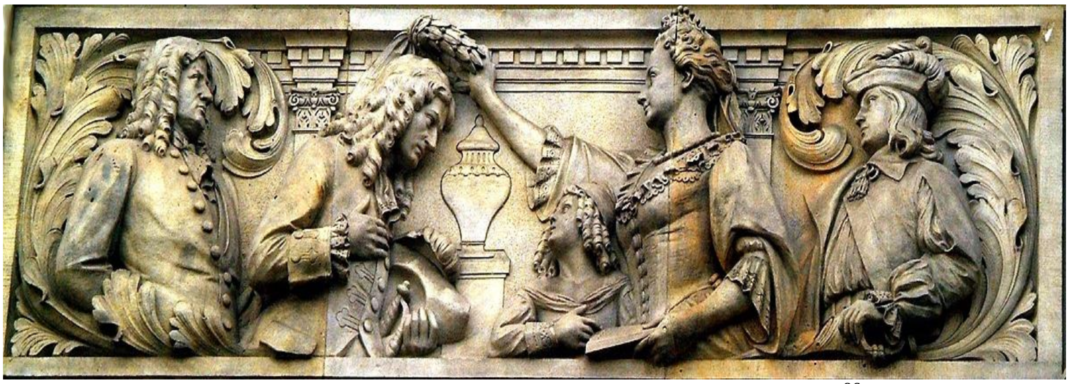


Figura 7: Sofia de Hanôver homenageando Leibniz com uma coroa de louros. [86]
Fonte: https://pt.wikipedia.org/wiki/Gottfried_Wilhelm_Leibniz (visitado em 01/01/2019)

---

[86] Sofia tornou-se amiga e admiradora de Gottfried Leibniz quando ele era bibliotecário na corte de Hanôver. A sua amizade durou desde 1676 até à sua morte em 1714. Esta amizade originou uma correspondência substancial publicada pela primeira vez no século XIX. **: https://pt.wikipedia.org/wiki/Gottfried_Wilhelm_Leibniz**

ANEXOS

ANEXO A

Conforme mencionamos em nossa introdução os diferentes tipos de "informações testáveis" geram problemas matemáticos diferentes. Também ressaltamos que as investigações sobre a fundamentação do MaxEnt podem contribuir para que futuramente possam se estabelecer novos princípios, a serem descobertos e que possibilitem, por exemplo, utilizar informações não-testáveis.

Neste anexo destacamos a abordagem de Jaynes sobre essas questões onde ele utiliza alguns exemplos para mostrar a distinção entre informações testáveis e não testáveis.

Aplicações e Recursos

Vamos estabelecer, para referência, um pouco do formalismo básico da Máxima Entropia para o caso discreto finito, deixando de fora generalizações até serem necessárias. Existem n diferentes possibilidades, que seriam distinguidas adequadamente por um único Índice (i = i, 2, ..., n). No entanto achamos útil, tanto para a notação quanto para as aplicações que temos em mente, introduzir, além disso, uma variável real x que pode assumir os valores discretos (x_i, $1< i \leq n$), definidos de alguma maneira e não, necessariamente, todos distintos. Se tivermos algumas informações sobre x, o problema é representá-las por meio de uma distribuição de probabilidade $\{p_i\}$ que tenha entropia máxima de acordo com I.

Claramente, esse problema não pode ser bem colocado para uma informação arbitrária; deve ser de tal forma que, dada a proposta de distribuição $\{p_i\}$, podemos determinar inequivocamente se devo ou não concorda com $\{p_i\}$. Essa informação será chamada <u>testável</u>. Por exemplo, considere:

I_1 = "é certo que tanh x < 0.7. "
I_2 = "Há pelo menos uma probabilidade de que 90% tanh x < 0.7".
I_3 = "O valor médio de tanh x é 0.675".
I_4 = "O valor médio de tanh x é provavelmente inferior a 0,7.
I_5 = "Há algum motivo para acreditar que tanhx = 0.675".

As declarações I_1, I_2, I_3 são testáveis e podem ser usadas como vínculos para maximizar a entropia. I_4 e I_5, embora seja claramente relevante para inferir sobre x, são muito vagas para serem testáveis e, atualmente, não temos um princípio formal porémpelo qual essas informações possam ser usadas em uma teoria matemática.

Entretanto, o nosso senso comum intuitivo faz uso de informações não testáveis e isso sugere que novos princípios devem ser descobertos para esses tipos de informações. (JAYNES 1978, p.240, grifo nosso, tradução nossa)[87]

[87] Let us set down, for reference, a bit of the basic Haximum Entropy formalism for the finite discrete case, putting off generalizations until they are needed. There are n different possibilities, which would be distinguished adequately by a single index (i = 1,2, ... ,n). Nevertheless we find it helpful, both for notation and for the applications we have in mind, to introduce in addition a real variable x, which can take on the discrete values (x_i, $1< i \leq n$), defined in any way and not necessarily all distinct. If we have certain information I about x, the problem is to represent this by a probability distribution {Pi} which has maximum entropy while agreeing with I. Clearly, such a problem cannot be well-posed for arbitrary information; I must be such that, given any proposed distribution {Pi}, we can determine unambiguously whether I does or does not agree with {Pi}. Such information will be called testable.
For example, consider:
I_I - "It is certain that tanh x < 0.7."
I_2 - "There is at least a 90% probability that tanh x < 0.7."
I_3- "The mean value of tanh x is 0.675."
I_4 - "The mean value of tanh x is probably less than 0.7."
I_5 - "There is some reason to believe that tanh x = 0.675."
Statements I_I, I_2, I_3 are testable, and may be used as constraints in maximizing the entropy. I_4 and I_5, although clearly relevant to inference about x, are too vague to be testable, and we have at present no formal principle by which such information can be used in a mathematical theory. However, the fact that our intuitive common sense does make use of nontestable information suggests that new principles for this, as yet undiscovered, must exist.

ANEXO B

Como ressaltamos no capítulo 4, Leibniz frequentemente enfatizava a necessidade de incluir a soma das séries no cálculo das probabilidades. Essa afirmação pode ser confirmada na seguinte correspondência que Leibniz enviou a Bernoulli em 1690.

RESPOSTA AO QUE O ILUSTRE J. BERNOULLI PUBLICOU EM MAIO DE 1690 NA ATA DOS ERUDITOS

Bernoulli me sugeriu um novo problema para resolver; que resolverei em breve, explicando o princípio da solução que ele trouxe para seu próprio problema, proposto no Jornal do qual falei. Eis a questão. Dois jogadores jogam com um único dado, de acordo com a seguinte regra: o vencedor será o primeiro a alcançar um número fixo de pontos com o dado. A começa jogando o dado uma vez, B também joga uma vez, depois A joga duas vezes, depois B também joga duas vezes, depois A joga três vezes e B três vezes, etc. Ou ainda: A começa jogando o dado uma vez, B então joga duas vezes, depois A joga três vezes, depois B quatro etc, até que uma dos dois vença. Estamos procurando as chances de cada um. Eu exemplico assim:

Seja $\frac{5}{6}$ = n, teremos $\frac{1}{6}$ = 1 - n. No primeiro caso temos:

1	n	n^2	n^3	n^4	n^5	n^6	n^7	n^8	n^9	n^{10}	n^{11}	etc.
A	B	A	A	B	B	A	A	A	B	B	B	etc

As chances de A são:

$1 + n + n^2 + n^3 + n^6 + n^7 + n^8 + n^{12} + n^{13} + n^{14} + n^{15}$ etc.

Após multiplicarmos por (1 – n), teremos:

$1 - n + n^2 - n^4 + n^6 - n^9 + n^{12} - n^{16}$ etc

Por outro lado as chances de B são:

$n + n^4 + n^5 + n^9 + n^{10} + n^{11} + n^{16} + n^{17} + n^{18} + n^{19}$ etc.

Após multiplicarmos por (1 – n), teremos:

$n - n^2 + n^4 – n^6 + n^9 - n^{12} + n^{16}$ etc.

No secundo caso teremos:

1 n n^2 n^3 n^4 n^5 n^6 n^7 n^8 n^9 n^{10} etc.
A B B A A A B B B B A etc

As chances de A são:

$1 + n^3 + n^4 + n^5 + n^{10} + n^{11} + n^{12} + n^{13} + n^{14}$ etc.

Após multiplicarmos pelo fator (1 - n), teremos:

$1 – n + n^3 – n^6 + n^{10} – n^{15}$ etc.

Por sua vez, as chances de B são:

$n + n^2 + n^6 + n^7 + n^8 + n^9 + n^{15} + n^{16} + n^{17}$ etc.

Após multiplicarmos pelo fator (1 - n), teremos:

$n – n^3 + n^6 – n^{10} + n^{15} – n^{21}$ etc.

Nos dois casos A + B = 1, a unidade que representa a totalidade da aposta. O mesmo método se aplica em casos semelhantes, onde haveria mais jogadores ou mais dados, deduzimos facilmente uma solução tão precisa quanto queremos. O problema é atraente porque, embora tenha uma aparência muito simples, leva a séries que ainda não estudamos o suficiente. (Leibniz, 1995, pp. 44-442, tradução nossa)[88]

[88] RÉPONSE À CE QUE L'ILLUSTRE J. BERNOULLI A PUBLIÉ AU MOIS DE MAI DE L'AN 1690 DANS LES ACTES DES ERUDITORUM.

Il m'a proposé un nouveau problem à résoudre; je vais y venir dans un instant, en expliquant le principe de la solution qu'il apportée à son prore problème, proposé dans le Journal dont j'ai parlé. Voici s'agit. Deux joueurs jouent avec un unique dé selon la règle suivant: le vainqueur sera le premier à réaliser avec le dé un nombre fixé de points. A commence et lance le dé une fois, B le lance une fois, puis A le lance deux fois, à la suite de quoi B le lance également deux fois, alors A lance trois fois et B trois fois, etc. Ou encore: A commence par lancer le dé une fois, B le lance alors deux fois, puis A trois fois, puis B quatre etc, jusqu'à ce que l'un des deux vainque. On cherche le rapport des chances. J'explique la chose ainsi: Soit 5/6 = n, on aura 1/6 = 1 - n. Dans le premier cas:
1 n n^2 n^3 n^4 n^5 n^6 n^7 n^8 n^9 n^{10} n^{11} etc.
A B A A B B A A A B B B etc. Les chances de A sont: $1 + n + n^2 + n^3 + n^6 + n^7 + n^8 + n^{12} + n^{13} + n^{14} + n^{15}$ etc. Après avoir effectué la multiplication par 1 – n nous aurons: $1 - n + n^2 - n^4 + n^6 - n^9 + n^{12} - n^{16}$ etc. Les chances de B sont en revanche: $n + n^4 + n^5 + n^9 + n^{10} + n^{11} + n^{16} + n^{17} + n^{18} + n^{19}$ etc. Multiplié par 1- n, achèvée nous obtiendrons: $n - n^2 + n^4 – n^6 + n^9 - n^{12} + n^{16}$ etc. Dans le second cas:
1 n n^2 n^3 n^4 n^5 n^6 n^7 n^8 n^9 n^{10} etc.
A B B A A A B B B B A etc. Les chances de A sont: $1 + n^3 + n^4 + n^5 + n^{10} + n^{11} + n^{12} + n^{13} + n^{14}$ etc. facteur de (1 - n) , c'est-à-dire après multiplication: $1 – n + n^3 – n^6 + n^{10} – n^{15}$ etc. Les chances de B sont à leur tour: $n + n^2 + n^6 + n^7 + n^8 + n^9 + n^{15} + n^{16}$ etc. Facteur de (1 - n) , soit quand on a effectué la multiplication: $n – n^3 + n^6 – n^{10} + n^{15} – n^{21}$ etc. Dans le deux cas A + B = 1, l'unité representant la totalité duu droit sur l'enjeu. La même Méthode vaut dans les cas similaires, où il y aurait davantage de joueurs ou davantage de dés, nous en déduisons aisément une solution aussi précise que nous voulons. Le probleme est séduisant car, quoique très simple en apparence, il conduit à des séries qu'on n'a pas encore etudiées de prés.