Instituto Federal de Educação Ciência e Tecnologia da Paraíba
Campus Campina Grande
Curso Superior de Licenciatura em Matemática

Sara Lopes Silva

O Teorema de Bayes: Aplicações e abordagens didáticas no ensino e na tomada de decisão

Campina Grande - PB
2025

Sara Lopes Silva

O Teorema de Bayes: Aplicações e abordagens didáticas no ensino e na tomada de decisão

Trabalho de Conclusão de Curso submetida à Coordenação do Curso de Licenciatura em Matemática, como parte dos requisitos para obtenção do título de graduado em Licenciatura em Matemática

Orientador: Prof. Me. Baldoino Sonildo da Nóbrega

Campina Grande - PB

2025

Catalogação na fonte:
Ficha catalográfica elaborada por Gustavo César Nogueira da Costa - CRB 15/479

S586t Silva, Sara Lopes
O Teorema de Bayes: aplicações e abordagens didáticas no ensino e na tomada de decisão / Sara Lopes Silva. . - Campina Grande, 2025.
51 f. : il.

Trabalho de Conclusão de Curso (Curso Superior de Licenciatura em Matemática) - Instituto Federal da Paraíba, 2025.
Orientador: Prof. Me. Baldoino Sonildo da Nóbrega.

1. Matemática 2. Teorema de Bayes 3. Ensino de matemática - probabilidade I. Nóbrega, Baldoino Sonildo da III. Título.

CDU519.2

Sara Lopes Silva

O Teorema de Bayes: Aplicações e abordagens didáticas no ensino e na tomada de decisão

Trabalho de Conclusão de Curso submetida à Coordenação do Curso de Licenciatura em Matemática, como parte dos requisitos para obtenção do título de graduado em Licenciatura em Matemática

Trabalho de conclusão de Curso. Campina Grande - PB, 14 de março de 2025:

Baldoino Sonildo da Nóbrega

Prof. Me. Baldoino Sonildo da Nóbrega
Orientador
IFPB

Prof. Dr. Jonathas Jerônimo Barbosa
Examinador Interno
IFPB

Alex Pereira Bezerra

Prof. Me. Alex Pereira Bezerra
Examinador Externo
UFRPE

Campina Grande - PB
2025

Agradecimentos

Agradeço primeiramente a Deus, por me conceder o dom da vida e por sua infinita bondade em cada passo da minha jornada. Aos meus pais e minhas irmãs, meu mais sincero agradecimento por todo amor, educação, apoio e incentivo. Cada conquista minha carrega um pouco de vocês, que sempre estiveram ao meu lado, respeitando minhas escolhas e me dando forças para seguir em frente. Nada disso seria possível sem os valores e ensinamentos que me transmitiram.

Aos meus professores do Instituto Federal da Paraíba, minha sincera gratidão por todo o conhecimento compartilhado e pelo apoio ao longo da minha jornada acadêmica. Em especial, ao meu orientador, Prof. Baldoino Sonildo da Nóbrega, agradeço pelas valiosas contribuições, paciência e dedicação durante o desenvolvimento deste trabalho.

“Se vi mais longe foi por estar de pé sobre os ombros de gigantes”

Resumo

A Teoria da Probabilidade desempenha um papel fundamental na análise de fenômenos aleatórios e na tomada de decisões em contextos de incerteza. Diante disso, este estudo tem como objetivo investigar as aplicações do Teorema de Bayes em diversos cenários, enfatizando sua utilização prática em áreas como saúde, negócios e educação matemática. No campo da saúde, o teorema contribui para a interpretação de diagnósticos médicos e a análise de testes clínicos. No setor financeiro, é aplicado na avaliação de riscos e na previsão de aprovações de empréstimos, auxiliando as instituições a determinar a probabilidade de inadimplência com base no histórico e no perfil de cada cliente. Já no contexto educacional, serão analisadas estratégias para o ensino do teorema, desde o nível básico até o superior. A proposta inclui uma abordagem didática que visa tornar o conceito mais acessível aos alunos, promovendo uma melhor compreensão da aplicação da probabilidade no cotidiano e em diferentes áreas do conhecimento. No entanto, o ensino desse conceito ainda enfrenta desafios devido à complexidade percebida por muitos estudantes. Este trabalho explora a importância da Probabilidade e do Teorema de Bayes, analisando suas aplicações práticas e ressaltando a necessidade de torná-los mais acessíveis no contexto educacional. A popularização desses conceitos pode contribuir para um melhor entendimento do raciocínio probabilístico e aprimorar a tomada de decisões baseadas em dados.

Palavras-chave: Probabilidade; Teorema de Bayes; Aplicações práticas

Abstract

The Theory of Probability plays a fundamental role in analyzing random phenomena and making decisions in uncertain contexts. In light of this, this study aims to investigate the applications of Bayes' Theorem in various scenarios, emphasizing its practical use in fields such as healthcare, business, and mathematical education. In the healthcare sector, the theorem contributes to the interpretation of medical diagnoses and the analysis of clinical tests. In the financial sector, it is applied to risk assessment and loan approval forecasting, helping institutions determine the probability of default based on each client's history and profile. In the educational context, strategies for teaching the theorem will be analyzed, from basic to higher education levels. The proposal includes a didactic approach designed to make the concept more accessible to students, promoting a better understanding of probability applications in daily life and across different fields of knowledge. However, teaching this concept still faces challenges due to the complexity perceived by many students. This study explores the importance of Probability and Bayes' Theorem, analyzing their practical applications and highlighting the need to make them more accessible in the educational context. Popularizing these concepts can contribute to a better understanding of probabilistic reasoning and enhance data-driven decision-making.

Keywords: Probability; Bayes' Theorem; Practical applications

Sumário

1 Introdução

A probabilidade é um dos pilares da Matemática, a qual tem um papel importante em diversas áreas do conhecimento. Sendo utilizada para compreender, modelar e analisar situações onde a incerteza está presente, desde a estatística básica até a inteligência artificial, ultrapassando áreas como a medicina e economia. A relevância da probabilidade está na capacidade de medir incertezas e oferecer métodos matemáticos que auxiliam na tomada de decisões, especialmente em situações quando não temos todas as informações. A teoria das probabilidades desempenha um papel essencial nessa interpretação de dados e na previsão de eventos futuros.

Diante disso, o Teorema de Bayes se destaca na análise de probabilidades condicionais, onde a partir de relações matemáticas é possível atualizar novos dados a probabilidade de um evento acontecer a medida que novas evidências são incorporadas ao calculo. As aplicações desse teorema são amplas e podem ser identificada em diversas áreas do conhecimento, como por exemplo no diagnóstico médico, onde auxilia na interpretação de testes clínicos, na segurança digital, onde é utilizado para identificar e-mails e na analise de crédito, fundamentando a aprendizagem sob o algoritmo de Naive Bayes, um dos métodos de classificação computacionalmente simples e prático.

Contudo, mesmo com o seu potencial e a sua abrangência, torna-se necessário que o ensino da probabilidade e do Teorema de Bayes seja mais acessível e de forma intuitiva. Deve-se abordar exemplos práticos e ferramentas visuais que possam facilitar e auxiliar na compreensão desses conceitos básicos. Isso vai permitir que os alunos desenvolvam um raciocínio lógico mais apurado, com o intuito de que compreendam como a probabilidade pode influenciar na tomada de decisões em diversas áreas do conhecimento. Com isso, o incentivo do conteúdo de probabilidade interligado com aplicações reais pode estimular o interesse dos estudantes, contribuindo para a disseminação desse conhecimento de forma mais ampla e eficaz.

Porém, existe alguns fatores que podem acabar dificultando o ensino do Teorema de Bayes na educação básica. Os alunos, nessa etapa, podem acabar considerando o teorema complexo de forma inicial, através de sua primeira interpretação, deduzindo ser abstrata e de difícil compreensão, especialmente para aqueles que não estão familiarizados com conceitos de probabilidade básica e de probabilidade condicional. Além disso, outro fato pode ser considerado como relevante para esse impasse, como por exemplo a sua exigência de raciocínio lógico, a qual nem sempre é desenvolvida no currículo tradicional, fazendo com o que muitos alunos observem o teorema de forma desafiadora e puramente teórico.

A fim de superar esses impasses, a utilização de aplicações práticas surge como uma estratégia para facilitar a compreensão do Teorema de Bayes, a partir de abordagens interativas onde a compreensão se torna mais intuitiva e eficaz, conectando o aprendizado teórico com exemplos práticos da realidade que o discente está inserido, com o foco de estimular o interesse e a demonstração da relevância desse conceito para a tomada de decisões, preparando cada discente para lidar com problemas reais que envolvem incerteza e a análise de dados.

1.1 Objetivo Geral

Explorar as aplicações práticas do Teorema de Bayes em diferentes contextos, como saúde, negócios, e inteligência artificial, avaliando sua eficácia na resolução de problemas, como também a sua presença em questões que envolva o raciocínio lógico.

1.2 Objetivos Específicos

- Identificar aplicações do Teorema de Bayes na área da saúde, no ramo dos negócios e no raciocínio lógico;

- Examinar contribuições do Teorema de Bayes para a inteligência artificial, abordando como ele é empregado em sistemas de aprendizado de máquina;

- Avaliar a eficiência do Teorema na resolução de problemas práticos;

2 Referencial Teórico

2.1 Probabilidade

2.1.1 Um pouco de História de Probabilidade

A composição em cima da teoria da probabilidade iniciou-se em meados do século XV, sendo como incentivo as apostas e jogos de azar, retratado a probabilidade de forma intuitiva. Existem registros sobre os primeiros indícios probabilísticos, a parir do jogo Tali, onde era utilizado pedaços de ossos de animais formando faces como as de um dado irregular. Desde então foram iniciados diversos outros registros matemáticos sobre os conceitos probabilísticos, a partir do século XVI, o italiano Girolamo Cardano (1501-1576), o qual é considerado o iniciador da teoria das probabilidades, por apresentar relações por meio da razão entre a combinatória dos casos possíveis de um evento ocorrer e o número de casos favoráveis dado o número de casos possíveis.

Um grande incentivo para o avanço da teoria sobre a probabilidade foi em 1657, com o holandês Christiaan Huygens (1629-1695), com a publicação da primeira pesquisa sobre o cálculo de probabilidade. Porém, apenas em 1713 que foi publicado primeiro livro sobre a teoria das probabilidades, escrito por Jakob Bernoulli (1655-1705), se referindo tanto aos jogos de azar quanto ao teorema de Bernoulli sobre as distribuições binomiais.

Já no século XVIII, o francês Thomas Bayes (1702-1761), iniciou uma outra nova abordagem de análise em probabilidade, denominada teorema das probabilidades condicionais, ou apenas Teorema de Bayes, que "consiste em determinar a probabilidade dos acontecimentos perante certas condições iniciais" (JUNQUEIRA, 2015, p. 4). Evidenciando duas noções de probabilidade, conhecidas como Objetiva e Subjetiva.

No entanto, o matemático francês Pierre Simon Laplace (1749-1827), foi quem introduziu oficialmente a probabilidade no mundo da "Matemática", em seu livro Théorie Analytique des Probabilités (Teoria Analítica das Probabilidades) de 1812. Segundo Junqueira (2015), a teoria das probabilidades tornou-se um eficaz instrumento, exato e fiável do conhecimento, nesse mesmo período, Andrei N. Kolmogorov (1903-1987), "que axiomatizou corretamente a Teoria das Probabilidades e um dos sucessos da sua abordagem foi dar uma definição rigorosa da expectância condicional" (JUNQUEIRA, 2015, p. 04). "A partir de Kolmogorov o desenvolvimento das probabilidades tem seu crescimento exponencial sendo hoje um ramo importante da matemática" (ALMEIDA, 2005, p. 15).

2.1.2 Probabilidade - Conceitos Básicos

A noção intuitiva de probabilidade é utilizada em diversos momentos do dia a dia, muitas vezes de forma imperceptível. Por exemplo, ninguém afirmaria que a probabilidade de um bebê nascer com dentes é exatamente zero. Dessa forma, a probabilidade permite calcular as chances de ocorrência de um determinado evento e avaliar a possibilidade de sua realização, conforme afirmam Bussab e Morettin (2010).

De acordo com Dantas (2013, p. 16), experimentos que, quando repetidos sob as mesmas condições, não produzem sempre o mesmo resultado são chamados de experimentos aleatórios. Um exemplo disso é o sorteio na loteria: os números são escolhidos de maneira aleatória, o que impede a previsão exata dos vencedores. Dessa forma, a chance de ganhar depende da quantidade de combinações realizadas. Além disso, o resultado final desse experimento pode ou não ser conhecido antecipadamente.

Experimento Aleatório

Definição 2.1.0 Experimentos que ao serem repetidos nas mesmas condições não produzem o mesmo resultado são denominados experimentos aleatórios.

Exemplo 2.1.0 Ao lançar uma moeda para o alto, podemos observar se a face que ficou para cima, ou não.

Definição 2.1.1 Os experimentos que ao serem repetidos nas mesmas condições conduzem ao mesmo resultado são denominados determinísticos.

Espaço Amostral

O espaço amostral é o conjunto que inclui todos os resultados obtidos do experimento aleatório.

Definição 2.1.2 Denominaremos espaço amostral associado a um experimento o conjunto de seus resultados possíveis.

O espaço amostral será representado por um conjunto S. Os elementos desse espaço serão conhecidos como eventos simples ou pontos amostrais. Quando realizamos um experimento aleatório, o espaço amostral será o conjunto de todos os resultados que podem ocorrer.

Quando o espaço amostral é finito ou infinito enumerável, é chamado espaço amostral discreto. Caso contrário, isto é, quando S é não enumerável, vamos chamá-lo de espaço amostral contínuo.

Exemplo 2.1.1 Três peças são retiradas de uma linha de produção. Cada peça é classificada em boa (B) ou em defeituosas (D). O espaço amostral associado a esse experimento é:

$$S = \{BBB, BBD, BDB, BDD, DBB, DBD, DDB, DDD\}.$$

Note que, no exemplo anterior o espaço amostral é finito. Exibiremos a seguir um exemplo onde o seu espaço amostral seja infinito.

Exemplo 2.1.2 Considere o experimento que consiste em retirar uma lâmpada de um lote e medir o seu "tempo de vida"antes de se queimar. Um espaço amostral conveniente é

$$S = \{t \in \mathbb{R} : t \geq 0\}$$

Isto é, o conjunto de todos os números reais não negativos. Se A indicar o evento "o tempo de vida da lâmpada é inferior a 20 horas", então A $= \{t : 0 \leq t \leq 20\}$. Esse é um exemplo de um espaço amostral *contínuo*.

Evento

Definição 2.1.3 Denominaremos de evento a todo resultado ou subconjunto de resultados de um experimento.

Os eventos serão representados por subconjuntos do espaço amostral, onde, será representado por um conjunto unitário. Logo, que envolvem apenas um ponto no espaço amostral, são conhecidos como eventos simples. O evento A, só vai ocorrer quando o resultado do experimento é um evento simples e vai pertencer a A (Dantas, 2013, p.19).

Exemplo 2.1.3 No exemplo anterior consideremos o evento A: duas peças são boas. Tem-se:

$$A = \{BBD, BDB, DBB\}$$

Então A ocorre se ocorrer um dos três eventos simples BBD, BDB, ou DBB.

2.1.3 Operações entre eventos

Definição 2.1.4 *A reunião de dois eventos A e B, denotada de $A \cup B$, é o evento que ocorre se pelo menos um deles ocorre.*

Definição 2.1.5 *A interseção de dois eventos A e B, denotada $A \cap B$, é o evento que ocorre se ambos ocorrem.*

Definição 2.1.6 *O complementar do evento A, denotado A^c, é o evento que ocorre quando A não ocorre.*

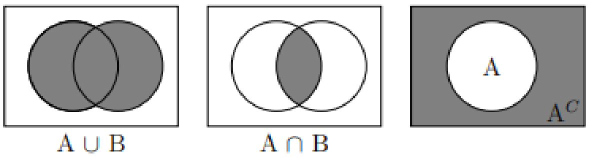

Figura 1 – União, interseção e complementar de eventos
Fonte: Dantas (2013)

2.1.4 Definição Clássica

Definição 2.1.7 Consideremos um espaço amostral S com N eventos simples, que suporemos igualmente possíveis. Seja A um evento de S composto de m eventos simples. A probabilidade de A, que denotaremos P(A), é definida por:

$$P(A) = \frac{m}{N}$$

Essa foi a primeira definição formal de probabilidade, a qual foi apresentada por Girolamo Cardano (1501-1576). E vamos nos referir a ela como a definição clássica de probabilidade. Vale destacar que essa definição se apoia em duas suposições principais:

1. Há um número finito de eventos elementares, isto é, S é um conjunto finito;
2. Os eventos elementares são igualmente prováveis.

Embora essas suposições restrinjam um pouco a aplicação da definição, ela ainda é muito significativa. Ao longo dos estudos, veremos que muitos problemas podem ser resolvidos com base nela, e sua compreensão será essencial para avançarmos em outros conceitos.

Exemplo 2.1.4 No lançamento de um dado, qual é a probabilidade de se obter face maior que 4?

Solução:

Sabemos que $n(S) = 6$ e também que o evento de interesse é A = $\{5, 6\}$.

Logo, P(A) = $\frac{2}{6}$

P(A) = $\frac{1}{3}$

2.1.5 Definição axiomática de probabilidade

Observamos que a definição clássica de probabilidade se restringe a espaços amostrais finitos, onde os eventos elementares são equiprováveis. Agora vamos analisar algo mais abrangente: a *definição axiomática de probabilidade*.

De acordo com a definição clássica, a probabilidade é um número atribuído a cada evento dentro do seu espaço amostral S. Logo, esse número, denominado probabilidade, possui propriedades específicas. A definição que apresentaremos considera os seguintes aspectos.

Definição 2.1.8 ***Definição axiomática de probabilidade***

Seja S um espaço amostral associado a um experimento aleatório. Probabilidade é uma função, denotada por P, que associa a cada evento A de S um número real P(A) que satisfaz os seguintes axiomas:

$$\textit{Axioma 1 : } P(A) \geq 0$$
$$\textit{Axioma 2 : } P(S) = 1$$
$$\textit{Axioma 3 : } A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$$

É importante que você observe que os três axiomas correspondem às três primeiras propriedades vistas para a definição clássica de probabilidade. Para a definição clássica, a demonstração da validade dessas três propriedades é imediata a partir da teoria de conjuntos. No caso geral, elas formam o conjunto de axiomas da probabilidade. Como todas as outras propriedades foram deduzidas a partir dessas três propriedades, elas continuam valendo no caso geral, ou seja, a partir dos três axiomas deduzimos as seguintes propriedades:

$$\mathrm{P}(\emptyset) = 0$$
$$\mathrm{P}(\overline{A}) = 1 - P(A)$$
$$\mathrm{P(A - B) = P(A) - P}(A \cap B)$$
$$\mathrm{P(B - A) = P(B) - P}(A \cap B)$$
$$\mathrm{P}(A \cup B) = \mathrm{P(A) + P(B) - P}(A \cap B)$$
$$A \subset B \Rightarrow P(A) \leq P(B)$$
$$\mathrm{P(A)} \leq 1$$

Exemplo 2.1.5 Dados S = {1, 2, 3} , A = {1} , B = {2} , C = {3} , P(A) = $\frac{1}{3}$, P(B) = $\frac{1}{3}$, calcule:

1. P(C)
2. P($A \cup B$)
3. P($\overline{A}$)
4. P($\overline{A} \cap \overline{B}$)

5. P($\overline{A} \cup \overline{B}$)

Solução:

1. Como $P(S) = 1$, resulta que $P(C) = 1 - P(A) - P(B) = \frac{1}{3}$.
2. Como A e B são mutuamente exclusivos, P($A \cup B$) = P(A) + P(B) = $\frac{2}{3}$.
3. P($\overline{A}$) =$1 - P(A)$ = $\frac{2}{3}$.
4. Pela lei de De Morgan, temos que P($\overline{A} \cap \overline{B}$) = P($\overline{A \cup B}$) =$1 - P$(A $\cup B$)=1- $\frac{2}{3} = \frac{1}{3}$.
5. Pela lei de De Morgan, temos que P($\overline{A} \cup \overline{B}$) = P($\overline{A \cap B}$) = $1 - P$(A $\cap B$)

$$= 1 - 0 = 1.$$

2.1.6 Probabilidade Condicional

Probabilidade condicional é a probabilidade de ocorrência de um evento A, dado que outro evento B já ocorreu. Matematicamente, é expressa pela fórmula:.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad \text{onde } P(B) > 0$$

Em outras palavras, a probabilidade condicional mede o quanto o conhecimento prévio da ocorrência do evento B influencia na probabilidade do evento A. Por exemplo, a probabilidade de chover amanhã pode mudar se soubermos que hoje está nublado, caracterizando, assim, uma probabilidade condicionada por essa nova informação.

2.1.7 Regra do Produto

A definição de probabilidade condicional leva a um resultado importante, conhecido como regra da multiplicação.

Proposição 2.1 Regra da multiplicação para 2 eventos

Sejam A e B eventos de um espaço amostral S. Então

$$\mathrm{P}(A \cap B) = \begin{cases} \mathrm{P}(B)\,\mathrm{P}(A \mid B) \\ \mathrm{P}(A)\,\mathrm{P}(B \mid A) \end{cases} \tag{2.1}$$

Esse resultado nos permite calcular a probabilidade da interseção de dois eventos e é muito útil para modelar experimentos que têm caráter sequencial, isto é, que são executados em etapas, uma seguida da outra.

2.1.8 Independência de Eventos

Definição 2.1.9 Sejam A e B eventos de um espaço amostral. Então, A e B são independentes se,

$$\mathrm{P}(A \mid B) = \mathrm{P}(A) \tag{2.2}$$

Essa definição traz consigo algumas implicações significativas. A primeira delas é:

$$\mathrm{P}(A \mid B) = \mathrm{P}(A) \Rightarrow \mathrm{P}(B \mid A) = \mathrm{P}(B)$$

De fato:

$$\begin{aligned} \mathrm{P}(A \mid B) = \mathrm{P}(A) \Rightarrow \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \mathrm{P}(A) \\ \Rightarrow \mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B) \Rightarrow \\ \mathrm{P}(B \mid A) = \frac{\mathrm{P}(B \cap A)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A)\,\mathrm{P}(B)}{\mathrm{P}(A)} = \mathrm{P}(B) \end{aligned}$$

Diante disso, concluímos que: Se A e B são independentes, então B e A também o são (comutatividade).

A segunda implicação, é a seguinte: Se A e B são independentes, então $\mathrm{P}(A \cap B) = \mathrm{P}\,(A)\,\mathrm{P}\,(B)$.

Logo, a recíproca dessa afirmativa também é verdadeira, ou seja, se:

$$\mathrm{P}(A \cap B) = \mathrm{P}\,(A)\,\mathrm{P}\,(B) \tag{2.3}$$

Sendo então A e B independentes. De fato:

$$\begin{aligned} \mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B) \Rightarrow \\ \mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(A)\,\mathrm{P}(B)}{\mathrm{P}(B)} = \mathrm{P}(A) \Rightarrow \\ \text{A e B são independentes.} \end{aligned}$$

Esse resultado nos permite estabelecer uma outra definição equivalente para a independência de dois eventos.

Definição 2.1.10 Sejam A e B eventos de um espaço amostral S. Então, A e B são independentes se:

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B) \tag{2.4}$$

Exemplo 2.1.5 A partir dos dados apresentados na tabela a seguir, referentes à participação de funcionários de uma empresa em planos de aposentadoria complementar:

		Plano pessoal		Total
		Sim	Nao	
Plano da	Sim	200	200	400
Empresa	Não	0	100	100
Total		200	300	500

Estudaremos os eventos E = "empregado tem o plano aposentadoria complementar da empresa" e por P = "empregado possui plano pessoal de aposentadoria complementar". Vamos ver se esses eventos são independentes.

Solução:

Temos que

$$\mathrm{P}(P) = \tfrac{2}{5}$$
$$\mathrm{P}(E) = \tfrac{4}{5}$$
$$\mathrm{P}(E \cap P) = \tfrac{2}{5} \neq \mathrm{P}(P)\,\mathrm{P}(E)$$

Sendo assim, os eventos P e E não são independentes. Outra forma de ver isso é

$$\mathrm{P}(E \mid P) = \tfrac{200}{200} = 1 \neq \mathrm{P}(E) = \tfrac{4}{5}$$

2.2 Teorema de Bayes

2.2.1 Thomas Bayes

Thomas Bayes (1702–1761), nascido na Inglaterra, em uma família de religiosos dissidentes, Bayes acabou estudando teologia e atuando como ministro na paróquia de Tunbridge Wells. Porém, finalizou os seus estudos de Teologia na Universidade de Edimburgo, na Escócia. No entanto, possuía um grande interesse por matemática e estatística. O passado de Bayes alternou entre a teologia e a Matemática, visto que, pouco se sabe sobre sua vida, mas por ser considerado como não-conformista, por ser membro da Igreja Presbiteriana, onde não era permitido se matricular em universidades locais (GELMAN et al. 2013).

Logo após a sua morte, o amigo e filósofo Richard Price (1723-1791), reconheceu a importância de suas ideias e apresentou um de seus trabalhos a Royal Society, uma sociedade científica britânica, exprimindo a sua teoria de probabilidade, com o nome de "Ensaio para resolver um problema na doutrina das probabilidades". Esse artigo possuía a demonstração do que agora é conhecido como o Teorema de Bayes, proporcionando uma probabilidade que não trabalha apenas com uma unidade fixa, mas com uma unidade que se evolui a partir de novas informações que vão sendo obtidas. No entanto, após a publicação de seu trabalho, o mesmo acabou caindo em esquecimento, que só foi resgatado de volta com o matemático francês Pierre-Simon de Laplace (1749-1827), que concluiu revelando ao mundo (GELMAN et al. 2013)

Thomas Bayes revolucionou o pensamento probabilístico ao fornecer uma maneira formal e coerente de atualizar probabilidades com base em novas informações. Sua ideia, inicialmente discreta, tornou-se uma das ferramentas mais importantes na ciência contemporânea. O Teorema de Bayes passou a ser essencial em áreas como a estatística, inteligência artificial e o aprendizado de máquina, o qual possui o intuito de estabelecer uma relação entre probabilidades condicionais, através da atualização de crenças e com base em novas evidências a cada resultado.

2.2.2 Teorema de Bayes

Sejam A (hipótese) e B (evidência) dois eventos em um espaço amostral, com $P(B) > 0$. O Teorema de Bayes afirma que:

$$P(A \mid B) = \frac{P(B \mid A).P(A)}{P(B)}. \tag{2.5}$$

Lema 2.2.0 Lei da Probabilidade Total (LPT)

Sejam $H_1, H_2, \ldots, H_n$ eventos que formam uma *partição* do espaço amostral S, isto é,

- $H_i \cap H_j = \varnothing$ para $i \neq j$,
- $H_1 \cup H_2 \cup \cdots \cup H_n = S$.

Para qualquer evento B, a Lei da Probabilidade Total diz que:

$$P(B) = \sum_{i=1}^{n} P(B \mid H_i).P(H_i). \tag{2.6}$$

Demonstração Formal do Teorema de Bayes

Definições e Notação

- A: evento (hipótese) cuja probabilidade queremos atualizar à luz de B.
- B: evidência observada.
- $P(A)$: probabilidade a priori de A (antes de observar B).
- $P(A \mid B)$: probabilidade a posteriori (depois de observar B).
- $P(B \mid A)$: verossimilhança (probabilidade de B, supondo que A seja verdadeiro).
- $P(B)$: probabilidade marginal de B, isto é, $P(B) = P(A \cap B) + P(A^c \cap B)$ ou, de forma geral, via partições.

Definição de Probabilidade Condicional

Por definição, se $P(B) > 0$, então

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}. \tag{2.7}$$

Logo, o problema se reduz a expressar $P(A \cap B)$ de uma forma conveniente.

Expressando $P(A \cap B)$

Também pela definição de probabilidade condicional, mas agora invertendo os papéis de A e B:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \quad \Longrightarrow \quad P(A \cap B) = P(B \mid A).P(A). \tag{2.8}$$

Substituição na Fórmula de Probabilidade Condicional

Substituindo $P(A \cap B) = P(B \mid A).P(A)$ em $P(A \mid B) = P(A \cap B)/P(B)$, obtemos:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \mid A).P(A)}{P(B)}.$$

Isto prova a Equação (2.5):

Cálculo de $P(B)$ **via Lei da Probabilidade Total**

Em muitos casos, $P(B)$ é decomposto pela Equação (2.6). Por exemplo, se A faz parte de uma partição $H_1, \ldots, H_5$:

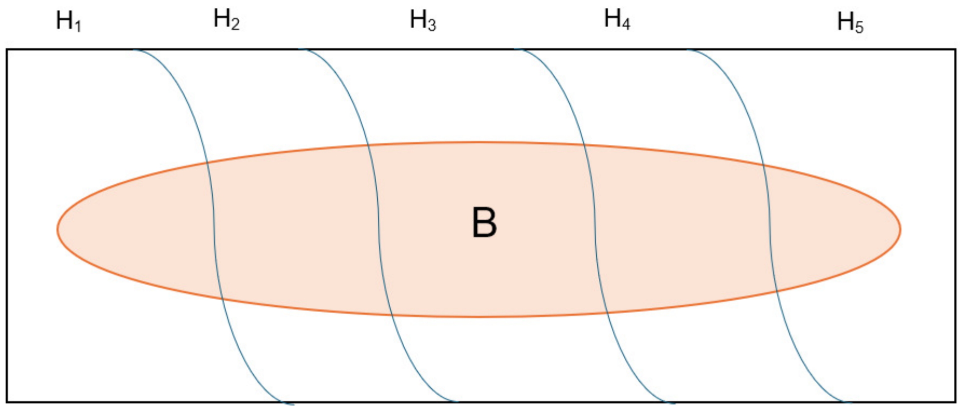

Figura 2 – Particionando um evento em vários subconjuntos mutuamente excludentes

Para qualquer evento B podemos escrever

$$B = (B \cap H_1) \cup (B \cap H_2) \cup (B \cap H_3) \cup (B \cap H_4) \cup (B \cap H_5)$$

Como $(B \cap H_1)$, $(B \cap H_2)$, $(B \cap H_3)$, $(B \cap H_4)$ e $(B \cap H_5)$ são mutuamente excludentes entre si. Pelo axioma (iii) da definição de probabilidade, temos:

$$P(B) = P\,(B \cap H_1) + P\,(B \cap H_2) + P\,(B \cap H_3) + P\,(B \cap H_4) + P\,(B \cap H_5)$$

Pelo teorema do Produto, temos

$$P\,(B \cap H_i) = P(H_i)P(B \mid H_i)$$

Dessa Forma, temos:

$$P(B) = P(H_1)P(B \mid H_1) + P(H_2)P(B \mid H_2) + P(H_3)P(B \mid H_3) + P(H_4)P(B \mid H_4) + P(H_5)P(B \mid H_5)$$

De forma geral, temos:

$$P(B) = \sum_{i=1}^{n} P(B \mid H_i).P(H_i). \tag{2.9}$$

Consequentemente, o Teorema de Bayes pode ser reescrito como:

$$P(H_k \mid B) = \frac{P(B \mid H_k), P(H_k)}{\sum_{i=1}^{n} P(B \mid H_i), P(H_i)}, \tag{2.10}$$

sendo H_k o evento de interesse (ou hipótese) dentro da partição.

O Teorema de Bayes é uma consequência natural da definição de probabilidade condicional e da Lei da Probabilidade Total. Ele fornece uma forma rigorosa de atualizar a probabilidade de um evento (ou hipótese) à medida que novas evidências (ou dados) são obtidas. Essa capacidade de *aprendizagem* torna o teorema fundamental em áreas como estatística, inteligência artificial, tomadas de decisão e muitas outras.

Exemplo 2.2.0 Em uma turma de Administração, 65% dos alunos são do sexo masculino. Sabe-se que 30% dos alunos tem carro, enquanto que essa proporção entre as alunas se reduz para 18%. Sorteia-se ao acaso um estudante dessa turma usando o seu número de matrícula e constata-se que possui um carro. Qual é a probabilidade de que a pessoa sorteada seja do sexo feminino?

Solução:

Os eventos em questão envolvem o sexo do aluno e a posse de um carro. Vamos definir os eventos de interesse da seguinte forma:

H = homem

M = mulher

C= possui carro

NC = não possui carro

Note que H e M definem uma partição do espaço amostral, assim como C e NC . No entanto, as probabilidades a priori dadas referem-se a H e M; logo, a partição de S será definida em termos desses eventos. Os dados do problema nos dão que

$$\mathrm{P(H)} = 0{,}65 \Rightarrow P(M) = 0,35$$

$$\mathrm{P(C} \mid H) = 0,30 \Rightarrow P(NC \mid H) = 0,70$$

$$\mathrm{P(C} \mid M) = 0,18 \Rightarrow P(NC \mid M) = 0,82$$

O problema pede P(M|C) e para calcular essa probabilidade, temos que calcular P(C). Pelo teorema da probabilidade total, sabemos que

$$\mathrm{P(C)} = \mathrm{P(C} \cap M) + P(C \cap H)$$

$$\mathrm{P(C)} = \mathrm{P(M)}\ \mathrm{P(C} \mid M) + P(H)P(C \mid H)$$

$$\mathrm{P(C)} = 0{,}35 \cdot 0,18 + 0,65 \cdot 0,30$$

$$\mathrm{P(C)} = 0{,}258$$

Logo,

$$P(M \mid C) = \frac{P(C \cap M)}{P(C)}$$

$$\text{P(M} \mid C) = \frac{P(M) \cdot P(C|M)}{P(C)}$$

$$\text{P(M} \mid C) = \frac{0{,}35 \cdot 0{,}18}{0{,}0{,}258}$$

$$\text{P(M} \mid C) = 0,24418$$

2.2.3 Classificador Naive Bayes

O algoritmo Naive Bayes é um dos métodos de aprendizado de máquina supervisionado (machine learning), desempenhando um papel fundamental dentro da Inteligência Artificial (IA). Geralmente, esse algoritmo é utilizado em problemas que requerem a classificação de dados, como no reconhecimento de padrões, diagnóstico médico, análise de sentimentos e detecção de spam. Ele permite o cálculo preciso de uma determinada classe com base em um conjunto de características ou evidências.

Sua base matemática está no Teorema de Bayes, desenvolvido por Thomas Bayes no século XVIII, o qual se fundamenta no uso de probabilidades condicionais. Esse modelo assume que as variáveis dos dados analisados são independentes entre si, o que facilita os cálculos e torna o algoritmo mais rápido e simples.

Sendo um dos algoritmos mais práticos e de simples compreensão, o Naive Bayes utiliza um classificador, o qual assume valores que são atribuídos condicionalmente e que são independentes entre si, devido a premissa da independência de cada evento ocorrer.

No contexto de classificação, queremos prever a classe C_k de um dado $X = (x_1, x_2, \ldots, x_n)$, ou seja, calcular:

$$P\left(C_k \mid X\right) = \frac{P\left(X \mid C_k\right) P\left(C_k\right)}{P(X)}$$

Pela suposição de independência dos atributos:

$$P\left(X \mid C_k\right) = P\left(x_1 \mid C_k\right) \cdot P\left(x_2 \mid C_k\right) \cdots P\left(x_n \mid C_k\right)$$

Assim, a fórmula final se torna:

$$P\left(C_k \mid X\right) = \frac{P\left(C_k\right) \prod_{i=1}^{n} P\left(x_i \mid C_k\right)}{P(X)}$$

Como $P(X)$ é constante para todas as classes, basta calcular:

$$P\left(C_k \mid X\right) = P\left(C_k\right) \prod_{i=1}^{n} P\left(x_i \mid C_k\right)$$

para cada classe C_k e escolher aquela com maior probabilidade.

O classificador funciona da seguinte forma:

Inicialmente, é realizada a coleta de dados para que ocorra o treinamento da máquina de acordo com o modelo a ser seguido. Em seguida, ocorre a preparação dos dados, etapa em que eles são pré-processados e submetidos a uma análise para sua normalização.

Após essa fase, passa-se à estimação das probabilidades, pois o algoritmo será capaz de calcular a probabilidade de ocorrência de cada classe ou categoria com base nos dados fornecidos na primeira etapa. Por fim, ocorre a classificação, responsável por categorizar os novos dados a partir das probabilidades estimadas anteriormente.

O Naive Bayes é um algoritmo de classificação baseado no Teorema de Bayes, amplamente utilizado devido à sua eficiência e simplicidade. Sua principal vantagem é a rapidez, pois exige pouco tempo de treinamento e lida bem com grandes volumes de dados. Além disso, é fácil de implementar, pois baseia-se apenas em estatísticas de contagem e multiplicação de probabilidades. Ele também apresenta ótimo desempenho para a classificação de textos, sendo bastante utilizado em filtragem de spam e análise de sentimentos. Outra vantagem é que, na maioria dos casos, não sofre com overfitting e pode funcionar bem mesmo com poucos dados.

Por outro lado, uma das principais limitações do Naive Bayes é a suposição de independência entre atributos, o que nem sempre reflete a realidade. Se as variáveis de entrada forem fortemente correlacionadas, o modelo pode apresentar baixa precisão. Além disso, ele enfrenta dificuldades quando há dados raros, pois se um atributo nunca apareceu na fase de treinamento, sua probabilidade será zero, o que inviabiliza certos cálculos – problema que pode ser corrigido com a suavização de Laplace. Outro ponto fraco é que o Naïve Bayes pode ter desempenho inferior a modelos mais sofisticados, como Árvores de Decisão e Redes Neurais, quando aplicado a conjuntos de dados complexos.

Em resumo, o Naive Bayes é uma excelente escolha quando se busca um algoritmo rápido, eficiente e fácil de implementar, especialmente para classificação de textos e problemas probabilísticos. No entanto, ele pode não ser a melhor opção quando os atributos são altamente correlacionados ou o problema exige um modelo mais complexo para alcançar alta precisão.

3 Materiais e Métodos

Para realizar esta pesquisa, empregamos uma metodologia de caráter exploratório, na qual foi conduzido uma revisão bibliográfica que teve como ponto inicial o acesso a materiais didáticos. Foram consultados livros, dissertações, teses e ampliado as buscas nos ambientes do Google Academic, Sciello e Portal da Capes. Além disso, foi utilizado artigos científicos especializados em aprendizado de máquina, que tratam da implementação do algoritmo Naive Bayes, foram cruciais para compreender como o teorema se traduz em soluções práticas de classificação e na tomada de decisão.

Inicialmente buscamos compreender os principais fundamentos sobre o Teorema de Bayes e algumas de suas aplicações e como ele é aplicado em contextos relacionados a saúde, negócios e segurança. Exploramos a utilidade do teorema ao ser aplicado de forma eficaz na resolução de problemas reais, além de destacar a importância de se tornar o ensino de probabilidade mais acessível e relevante para os estudantes.

4 Aplicações

4.1 Diagnóstico Médico

O Teorema de Bayes é amplamente aplicado em diversas áreas para atualizar probabilidades com base em novas evidências. Na medicina, auxilia no diagnóstico de doenças considerando a precisão dos testes. Em inteligência artificial, é usado em classificação de e-mails, análise de sentimentos e sistemas de recomendação. Além disso, aparece em finanças, ciência forense e engenharia, ajudando na tomada de decisões probabilísticas. Também é aplicado em problemas clássicos de probabilidade, como o Problema de Monty Hall, desafiando a intuição. A seguir veremos algumas aplicações desse teorema.

4.1.1 Aplicação 01: Diagnóstico de COVID-19

A pandemia de COVID-19, causada pelo vírus SARS-CoV-2, foi um dos maiores desafios sanitários da história recente. Surgiu no final de 2019 e rapidamente se espalhou pelo mundo, levando à adoção de medidas como distanciamento social, lockdowns e vacinação em massa. Durante esse período, os testes diagnósticos foram amplamente utilizados para detectar a infecção, especialmente testes rápidos de antígeno e exames de RT-PCR.

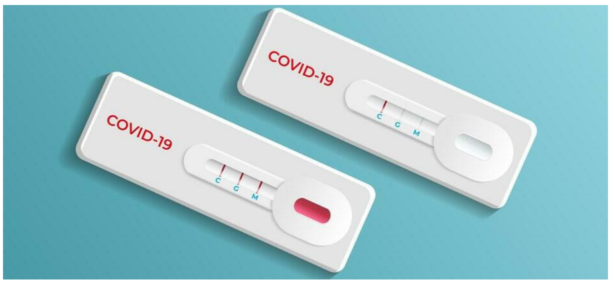

Figura 3 – Imagem de um teste de covid
(Fonte: coronavirus.saude.mg.gov.br)

Entretanto, nenhum teste é 100% preciso. Testes rápidos de antígeno, por exemplo, podem produzir falsos positivos (quando uma pessoa saudável recebe um resultado positivo) e falsos negativos (quando uma pessoa infectada recebe um resultado negativo). Isso levanta uma questão importante: dado um teste positivo, qual a real chance da pessoa estar infectada?

Considere o seguinte cenário:

Um teste rápido de antígeno para COVID-19 possui:

Sensibilidade (taxa de verdadeiro positivo): 85% (ou seja, identifica corretamente 85% dos infectados).

Especificidade (taxa de verdadeiro negativo): 98% (ou seja, identifica corretamente 98% dos não infectados).

A prevalência da COVID-19 na população testada é estimada em 2% (ou seja, apenas 2% das pessoas realmente têm COVID-19).

No dia 24 de junho de 2020, você realizou um teste de COVID-19 e o resultado foi POSITIVO. Utilizando o Teorema de Bayes, determine a probabilidade real de que você estava infectada dado que o teste foi positivo $P(COVID|TestePositivo)$.

Solução: Dados do Problema:

Foi considerado um teste rápido de antígeno para COVID-19 com as seguintes características:

- **Sensibilidade** ($P(\text{Positivo}|\text{COVID})$): 85% (0.85);
- **Especificidade** ($P(\text{Negativo}|\neg\text{COVID})$): 98% (0.98);
- **Prevalência da COVID-19** ($P(\text{COVID})$): 2% (0.02).

Nosso objetivo é determinar $P(\text{COVID}|\text{Positivo})$, ou seja, a probabilidade de que uma pessoa que testou positivo esteja realmente infectada.

Aplicação do Teorema de Bayes

O **Teorema de Bayes** é dado por:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{4.1}$$

onde:

- $P(A|B)$ é a probabilidade da pessoa estar infectada dado que o teste foi positivo;
- $P(B|A)$ é a sensibilidade do teste (0.85);
- $P(A)$ é a prevalência da doença na população (0.02);
- $P(B)$ é a probabilidade total de um teste positivo.

A probabilidade total de um teste positivo é dada por:

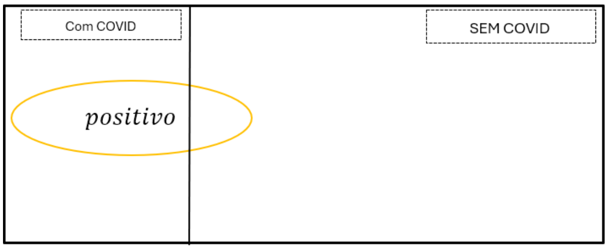

Figura 4 – Imagem do evento Positivo

$$Positivo = (Positivo \cap COVID) \cup (Positivo \cap SemCOVID)$$

$$P(\text{Positivo}) = P(\text{Positivo}|\text{COVID})P(\text{COVID}) + P(\text{Positivo}|\neg\text{COVID})P(\neg\text{COVID})$$

Sabemos que $P(\text{Positivo}|\neg\text{COVID})$ é a taxa de falsos positivos, dada por:

$$P(\text{Positivo}|\neg\text{COVID}) = 1 - \text{Especificidade} = 1 - 0.98 = 0.02$$

Além disso, $P(\neg\text{COVID})$ é a proporção da população que não está infectada:

$$P(\neg\text{COVID}) = 1 - P(\text{COVID}) = 1 - 0.02 = 0.98$$

Agora podemos calcular $P(\text{Positivo})$:

$$P(\text{Positivo}) = (0.85 \times 0.02) + (0.02 \times 0.98) = 0.017 + 0.0196 = 0.0366$$

Por fim, aplicamos o Teorema de Bayes:

$$P(\text{COVID}|\text{Positivo}) = \frac{P(\text{Positivo}|\text{COVID})P(\text{COVID})}{P(\text{Positivo})} = \frac{0.85 \times 0.02}{0.0366} = \frac{0.017}{0.0366} \approx 0.464$$

Conclusão

Apesar de o teste ter uma sensibilidade de 85% e uma especificidade de 98%, devido à baixa prevalência da doença na população (2%), a probabilidade de uma pessoa realmente estar infectada ao receber um teste positivo é de apenas **46,4%**.

Isso significa que, em uma população de baixa prevalência, um número significativo de testes positivos pode ser **falso positivo**. Dessa forma, um resultado positivo por si só

não deveria ser utilizado isoladamente para um diagnóstico definitivo de COVID-19. A melhor abordagem clínica seria combinar o resultado do teste com a avaliação de sintomas (febre, tosse seca persistente fadiga, perda de olfato/paladar etc) e outros exames confirmatórios, como o RT-PCR, que apresenta maior precisão e é considerado o padrão-ouro para o diagnóstico da infecção [**larremore2021test**, **cdc2021antigen**].

4.1.2 Aplicação 02: Câncer de Mama

Recomenda se que, a partir dos 40 anos, as mulheres façam mamografias anuais. Nessa idade, 1% (0,01) das mulheres são portadoras de um tumor assintomático de mama, ou seja, 99% (0,99) não tem câncer. Sabe-se que a mamografia apresenta resultado positivo em 80% dos casos. Os resultados dos teste indicam também que a mamografia apresenta resultado positivo em 80% das mulheres com câncer de mama, mas esse resultado também ocorre com 9,6% das mulheres sem o câncer, isto é, a mulher pode ter o resultado positivo mesmo sem ter propriamente câncer, denominamos isso de falso positivo. Diante desses dados uma mulher de 45 anos realizou o exame de mamografia anual, seguindo as recomendações do seu médico. O resultado do exame de mamografia foi positivo. Quais são as chances de realmente se ter câncer, dado que o teste deu positivo?

Solução:

Vamos analisar o caso e extrair as informações:

- A ocorrência do câncer de mama ser assintomático em mulheres dessa faixa etária é de 1% (0,01);
- A probabilidade de a mamografia apresentar um resultado positivo em mulheres com câncer de mama é de 80% (0,8);
- A probabilidade de a mamografia apresentar um resultado positivo em mulheres sem câncer de mama é de 9,6% (0,096).

Organizando as informações em uma tabela, temos:

#	Câncer (1%)	Sem câncer (99%)
Teste Positivo	80%	$9,6\%$
Teste Negativo	20%	$90,4\%$

Ao aplicarmos o Teorema de Bayes, veremos que:

$$P(\text{Câncer} \mid \text{Positivo}) = \frac{P(\text{Positivo} \mid \text{Câncer}) \cdot P(\text{Câncer})}{P(\text{Positivo})}$$

Onde:

- $P(Câncer|Positivo)$ é a probabilidade de a mulher ter câncer, visto que o resultado da mamografia foi positivo;
- $P(Positivo|Câncer)$ é a probabilidade de a mamografia dar positivo quando a mulher realmente tem câncer = 0,8;
- $P(Câncer)$ é a prevalência de câncer de mama na população feminina dessa faixa etária = 0,01;
- $P(Positivo|SemCâncer)$ é a probabilidade de a mamografia dar positivo quando a mulher não tem câncer = 0,096;
- $P(SemCâncer)$ é a probabilidade de a mulher não ter câncer = 0,99;
- $P(Positivo)$ é a probabilidade total do resultado positivo na mamografia.

Primeiro passo: Calcular $P(Positivo)$

Primeiramente será necessário efetuar o calculo da probabilidade total de se obter um resultado positivo para o teste de mamografia, através da soma das probabilidades do teste, ou seja:

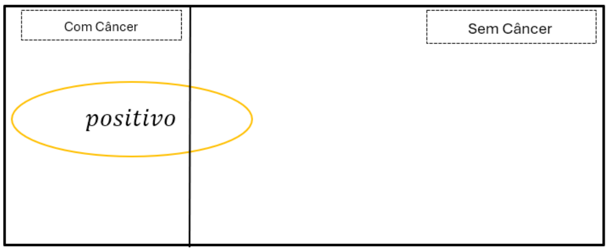

Figura 5 – Ilustração do evento positivo no conjunto universo

$$Positivo = (Positivo \cap Câncer) \cup (Positivo \cap SemCâncer)$$
$$P(Positivo) = P(Positivo \cap Câncer) + P(Positivo \cap SemCâncer)$$

$$P(Positivo) = P(Positivo|Câncer) \cdot P(Câncer) + P(Positivo|SemCâncer) P(SemCâncer)$$

Substituindo, obtemos:

$$P(Positivo) = 0,08 \cdot 0,01 + 0,096 \cdot 0,99$$
$$P(Positivo) = 0,008 + 0,09504$$
$$P(Positivo) = 0,10304$$

Logo, a probabilidade do teste dar positivo é de 10,3%

Segundo passo: Aplicando o Teorema de Bayes

A partir do resultado anterior, conseguimos aplicar o Teorema de Bayes, a fim de calcular a probabilidade de ter o câncer de mama, visto que o resultado da mamografia foi positivo:

$$P(Câncer|Positivo) = \frac{P(Positivo|Câncer) \cdot P(Câncer)}{P(Positivo)}$$

Substituindo os valores, obtemos:

$$P(Câncer|Positivo) = \frac{0{,}8 \cdot 0{,}01}{0{,}10304}$$

$$P(Câncer|Positivo) = \frac{0{,}008}{0{,}10304}$$

$$P(Câncer|Positivo) \approx 0,0776$$

Sendo assim, a probabilidade de realmente ter o câncer de mama, visto que o resultado da mamografia foi positivo é de aproximadamente 7,76%. Isso é relativamente baixo devido a prevalência do câncer de mama (1%) e a possibilidade de existência de falsos positivos. Observamos que a partir dos dados fornecidos na situação será importante sempre realizar um outro teste para confirmar o resultado positivo.

4.1.3 Aplicação 03: Testes de Gravidez

Os testes de gravidez (principalmente os de farmácia) medem a presença do hormônio hCG (gonadotrofina coriônica humana) na urina. Logo após a implantação embrionária, os níveis de hCG começam a subir, permitindo que o teste detecte a gravidez. Ainda assim, tais testes não são 100% perfeitos. Em termos numéricos, os testes comercializados no Brasil têm sensibilidade entre 97% e 99,5%

Sensibilidade é a probabilidade de o teste dar positivo se a mulher estiver realmente grávida (capacidade de detectar corretamente a gravidez) e especificidade é a probabilidade de o teste dar negativo se a mulher não estiver grávida (capacidade de detectar corretamente a ausência de gravidez).

Na prática, sempre há uma taxa de falso positivo (quando o teste dá positivo, mas a mulher não está grávida e uma taxa de falso negativo (quando o teste dá negativo, mas a mulher está grávida).

Observe o seguinte cenário:

Lica, uma mulher de 30 anos, está tentando engravidar há alguns meses. Após um atraso menstrual de cinco dias, decide realizar um teste de gravidez desses de farmácia, o qual obteve o resultado positivo para o teste. Sendo assim, Lica deseja saber qual é a chance dela está realmente gravida, uma vez que o resultado deu positivo? Suponha que a probabilidade de Lica estar grávida seja de 5% (0,05).

Figura 6 – Teste de gravidez

O teste de gravidez utilizado tinha as seguintes informações:

- Sensibilidade: 99% (probabilidade de teste positivo se a mulher está grávida).
- Especificidade: 99% (probabilidade de teste negativo se a mulher não está grávida), o que implica em 1% de falso positivo.

Solução:

Utilizamos o Teorema de Bayes para solucionar a probabilidade desse evento, dado o conjunto de evidências. Como queremos calcular a probabilidade de Lica está realmente gravida, dado que o resultado do teste de gravidez foi positivo.

A fórmula a partir do Teorema de Bayes será:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

onde:

- A é o evento "estar grávida".

- B é o evento "teste positivo".

No nosso caso, queremos $P(Grávida|Positivo)$.

Primeiro Passo: Calcular $P(Positivo)$:

$$P(Positivo) = P(Positivo \mid Grávida) \cdot P(Grávida) + P(Positivo \mid Nãográvida) \cdot P(Nãográvida)$$

Substituindo os valores:

$$P(\text{ Positivo }) = 0,99 \cdot 0,05 + 0,01 \cdot 0,95 = 0,0495 + 0,0095 = 0,059$$

Na situação formulada a probabilidade do teste dar positivo (considerando todos os casos) é de 5,9%.

Segundo Passo: Aplicando o Teorema de Bayes

Agora, iremos calcular P (Grávida | Positivo),

$$P(\text{ Grávida } \mid \text{ Positivo }) = \frac{P(\text{ Positivo } \mid \text{ Grávida })P(\text{ Grávida })}{P(\text{ Positivo })}$$

Substituindo:

$$P(\text{ Grávida } \mid \text{ Positivo }) = \frac{0,99 \times 0,05}{0,059} = \frac{0,0495}{0,059} \approx 0,838$$

Nesse caso, a probabilidade de Lica estar realmente grávida dado que o teste deu positivo é de aproximadamente 83,8%. Portanto, a probabilidade de estar grávida mesmo com o teste dando positivo ainda não é de 100%, devido a probabilidade a priori da gravidez que foi considerada é de apenas 5%, a qual é relativamente baixa e a probabilidade de existência do falso positivo (1%), acaba contribuindo para a probabilidade final desse teste.

Em situações reais, essa probabilidade a priori pode variar muito (por exemplo, se a mulher tem maior ou menor suspeita de gravidez, ou se o atraso menstrual é significativo, etc.). Além disso, resultados positivos geralmente são confirmados com um segundo teste ou com um exame de sangue quantitativo de HCG para maior certeza.

4.2 Inteligência Artificial e Aprendizado de Máquina

Classificação de E-mails: determinar se um e-mail é spam ou não com base em palavras-chave (filtros de spam). No auxílio para a tomada de decisões por meio de um classificador que decidirá se vai ser um bom ou mau negócio para o banco.

4.2.1 Aplicação 04: Classificação de E-mails

A empresa *SARA Empreendimentos LTDA* recebe diariamente inúmeros e-mails e deseja automatizar a filtragem das mensagens, classificando-as em **Spam** ou **Não Spam**. Para isso, optou por utilizar o **algoritmo Naive Bayes**, o qual é baseado no *Teorema de Bayes* e e na suposição de independência condicional entre as variáveis (no caso, as palavras presentes no e-mail).

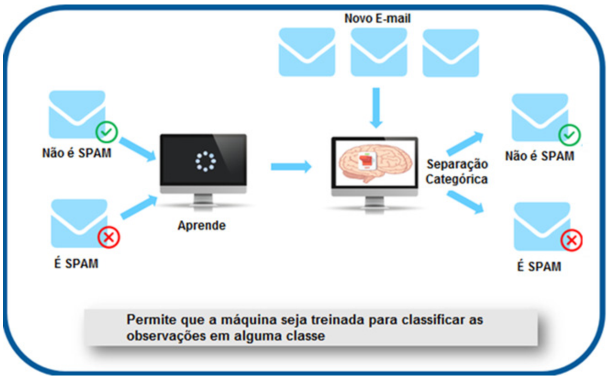

Figura 7 – Ilustração da classificação de e-mail

Contexto da Base de Treinamento:

1. A empresa coletou 100 e-mails rotulados:
   - 60 e-mails marcados como *Spam*.
   - 40 e-mails marcados como *Não Spam*.
2. A partir desses e-mails, extraiu-se a frequência das palavras que apareceram nos e-mails *Spam* e *Não Spam*, conforme Tabela de Frequências abaixo.
3. Um novo e-mail contendo as palavras *"clique no link"* precisa ser classificado.

Tabela de Frequências			
Palavra	**Spam = sim**	**Spam = não**	
clique	68	3	71
no	42	115	157
link	13	5	18
abaixo	7	1	8
confira	9	0	9
vamos	2	76	78
Total	**141**	**200**	**341**

Observação: O total na coluna $Spam = sim$ (141) representa a soma de ocorrências de todas as palavras em todos os e-mails rotulados como *Spam*. O mesmo vale para $Spam = não$ (200).

Probabilidades a priori

$$P(\text{Spam}) = \frac{60}{100} = 0{,}60, \quad P(\text{Não Spam}) = \frac{40}{100} = 0{,}40.$$

Probabilidades condicionais das palavras

Precisamos de:

$$P(\text{clique} \mid \text{Spam}) = \frac{68}{141} \approx 0{,}482,$$
$$P(\text{no} \mid \text{Spam}) = \frac{42}{141} \approx 0{,}298,$$
$$P(\text{link} \mid \text{Spam}) = \frac{13}{141} \approx 0{,}092,$$

e

$$P(\text{clique} \mid \text{Não Spam}) = \frac{3}{200} = 0{,}015,$$
$$P(\text{no} \mid \text{Não Spam}) = \frac{115}{200} = 0{,}575,$$
$$P(\text{link} \mid \text{Não Spam}) = \frac{5}{200} = 0{,}025.$$

Chegada do novo e-mail "clique no link"

Seja $E = \{\textit{clique}, \textit{no}, \textit{link}\}$. No *Naive Bayes*, assumimos as palavras como independentes, então:

$$P(\text{Spam} \mid E) = P(\text{Spam}) \times P(\text{clique} \mid \text{Spam}) \times P(\text{no} \mid \text{Spam}) \times P(\text{link} \mid \text{Spam}),$$

$$= 0{,}60 \times 0{,}482 \times 0{,}298 \times 0{,}092 \approx 0, 0792$$

$$P(\text{Não Spam} \mid E) = P(\text{Não Spam}) \times P(\text{clique} \mid \text{Não Spam}) \times P(\text{no} \mid \text{Não Spam}) \times P(\text{link} \mid \text{Não Spam})$$

$$= 0{,}40 \times 0{,}015 \times 0{,}575 \times 0{,}025 = 0{,}00008625$$

Dessa forma observa-se que a classe de maior probabilidade

Conclusão

Com esses dados, o **Naive Bayes** classifica o novo e-mail "clique no link" como *Spam*, pois tem probabilidade maior para essa classe.

Observação Final:

- O Naive Bayes, ainda que faça a suposição simplificadora de independência entre as palavras, costuma funcionar muito bem em filtragem de mensagens, por ser computacionalmente eficiente e lidar bem com muitos recursos (tokens).

4.2.2 Aplicação 05: Tomada de Decisão sobre Empréstimos

Maria sempre teve o sonho de abrir seu próprio negócio e depois de uma conversa com seus amigos e familiares, decidiu que já estava na hora de seguir com ele e foi até a agência bancária mais próxima a fim de solicitar um empréstimo para abrir a sua pequena loja de roupas. Porém, antes de fazer a sua solicitação, Maria sabia que precisaria provar sua capacidade financeira ao banco. Ao chegar ao banco, o gerente explicou que o pedido de empréstimo passaria por uma análise de crédito, onde seriam consideradas diversas informações sobre o seu perfil financeiro. A concessão do empréstimo dependeria do resultado dessa análise.

O banco utiliza um **modelo estatístico de decisão baseado no Algoritmo Naive Bayes** para classificar clientes entre **"Aprovado"** e **"Negado"** para o empréstimo. Para isso, algumas informações são levadas em consideração:

- **Histórico com o banco**: Sem dívidas ou Crítico em Dívidas;
- **Estado Civil**: Solteiro ou Casado;
- **Tempo de Emprego**: Mais de 2 anos ou Menos de 2 anos;
- **Posse de Casa**: Casa própria ou Alugada.

Dados de Maria

Após fornecer suas informações ao gerente, os dados de Maria foram registrados da seguinte forma:

- **Histórico com o banco**: Sem dívidas;
- **Estado Civil**: Solteira;
- **Tempo de Emprego**: Pouco mais de um ano (<2 anos);
- **Posse de Casa**: Alugada.

Dado esse perfil, o banco utilizará a análise de **Naive Bayes** para determinar a probabilidade de Maria ter seu **empréstimo aprovado** ou **negado**. Para isso, considere a seguinte **tabela de treinamento** baseada no histórico de clientes anteriores:

Legenda:

- **Situação**: Gênero e estado civil do cliente.

Tabela 1 – Tabela de Treinamento para Análise de Crédito

Situação	**Histórico**	**Emprego**	**Casa**	**Risco**
H Solteiro	Existe Dívidas	$\leq$ 2 anos	Própria	Ruim
H Solteiro	Existe Dívidas	$\leq$ 2 anos	Alugada	Ruim
M Divorciada	Existe Dívidas	$\leq$ 2 anos	Própria	Bom
M Solteira	Sem Dívidas	$\leq$ 2 anos	Própria	Bom
M Solteira	Crítico Dívidas	$>$ 2 anos	Própria	Bom
M Solteira	Crítico Dívidas	$>$ 2 anos	Alugada	Ruim
M Divorciada	Crítico Dívidas	$>$ 2 anos	Alugada	Bom
H Solteiro	Sem Dívidas	$\leq$ 2 anos	Própria	Ruim
H Solteiro	Crítico Dívidas	$>$ 2 anos	Própria	Bom
M Solteira	Sem Dívidas	$>$ 2 anos	Própria	Bom
H Solteiro	Sem Dívidas	$>$ 2 anos	Alugada	Bom
M Divorciada	Sem Dívidas	$\leq$ 2 anos	Alugada	Bom
M Divorciada	Existe Dívidas	$>$ 2 anos	Própria	Bom
M Solteira	Sem Dívidas	$\leq$ 2 anos	Alugada	Ruim

- **Histórico**: Classificação do relacionamento financeiro do cliente com o banco (*Sem Dívidas*, *Crítico Dívidas*, *Existe Dívidas*).
- **Emprego**: Tempo de emprego do cliente, podendo ser superior a 2 anos ($>$ 2 anos) ou menor/igual a 2 anos ($\leq$ 2 anos).
- **Casa**: Se a pessoa mora em *Casa Própria* ou *Alugada*.
- **Risco**: Classificação de risco do cliente, podendo ser *Bom* ou *Ruim*.

Questões a serem avaliadas:

1. Usando o **Teorema de Bayes** e o Algoritmo **Naive Bayes**, calcule a probabilidade de Maria ter seu empréstimo aprovado.
2. Com base no cálculo anterior, a solicitação de Maria deve ser **Aprovada** ou **Negada**? Justifique sua resposta.
3. Explique como a hipótese de independência entre as variáveis influencia no resultado do modelo Naive Bayes.

Resolução: Análise de Crédito com Naive Bayes

Cálculo da Probabilidade de Aprovação do Empréstimo de Maria

A tabela de treinamento contém **14 registros** de clientes, com as seguintes classificações de risco:

- **Clientes com risco BOM (empréstimo aprovado)**: 9 registros.

- **Clientes com risco RUIM (empréstimo negado)**: 5 registros.

Dessa forma, a probabilidade a priori de cada classificação é:

$$P(\text{Bom}) = \frac{9}{14} = 0{,}6429, \quad P(\text{Ruim}) = \frac{5}{14} = 0{,}3571.$$

Agora, consideremos as informações de Maria:

- **Histórico com o banco**: Sem dívidas;
- **Estado Civil**: Solteira;
- **Tempo de Emprego**: ≤ 2 anos;
- **Posse de Casa**: Alugada.

Cálculo das Probabilidades Condicionais

Precisamos calcular as probabilidades condicionais de cada variável para as classes de **Bom** e **Ruim**. A partir da tabela de treinamento:

$$P(\text{Sem Dívidas} \mid \text{Bom}) = \frac{4}{9} \quad P(\text{Sem Dívidas} \mid \text{Ruim}) = \frac{2}{5}$$

$$P(\text{Solteira} \mid \text{Bom}) = \frac{3}{9} \quad P(\text{Solteira} \mid \text{Ruim}) = \frac{2}{5}$$

$$P(\leq 2 \text{ anos} \mid \text{Bom}) = \frac{3}{9} \quad P(\leq 2 \text{ anos} \mid \text{Ruim}) = \frac{4}{5}$$

$$P(\text{Alugada} \mid \text{Bom}) = \frac{3}{9} \quad P(\text{Alugada} \mid \text{Ruim}) = \frac{3}{5}$$

Aplicação do Teorema de Bayes

Para calcular $P(\text{Bom} \mid \text{Dados de Maria})$, utilizamos:

$$P(\text{Bom} \mid E) = \frac{P(E \mid \text{Bom})P(\text{Bom})}{P(E \mid \text{Bom})P(\text{Bom}) + P(E \mid \text{Ruim})P(\text{Ruim})}.$$

Passo 1: Cálculo de $P(E \mid \textbf{Bom})$ e $P(E \mid \textbf{Ruim})$

$$P(E \mid \text{Bom}) = P(\text{Sem Dívidas} \mid \text{Bom})P(\text{Solteira} \mid \text{Bom})P(\leq 2 \text{ anos} \mid \text{Bom})P(\text{Alugada} \mid \text{Bom})$$

$$= \frac{4}{9} \times \frac{3}{9} \times \frac{3}{9} \times \frac{3}{9}.$$

$$\approx 0{,}0164.$$

$$P(E \mid \text{Ruim}) = P(\text{Sem Dívidas} \mid \text{Ruim})P(\text{Solteira} \mid \text{Ruim})P(\leq 2 \text{ anos} \mid \text{Ruim})P(\text{Alugada} \mid \text{Ruim})$$

$$= \frac{2}{5} \times \frac{2}{5} \times \frac{4}{5} \times \frac{3}{5}.$$

$$\approx 0{,}0768.$$

Passo 2: Cálculo da probabilidade final

$$P(\text{Bom} \mid E) = \frac{(0{,}0164) \times (0{,}6429)}{(0{,}0164) \times (0{,}6429) + (0{,}0768) \times (0{,}3571)}.$$

$$P(\text{Bom} \mid E) = \frac{0{,}0105}{0{,}0105 + 0{,}0274} = \frac{0{,}0105}{0{,}0379} \approx 0{,}277.$$

Portanto,

$$P(\text{Ruim} \mid E) = 1 - P(\text{Bom} \mid E) = 1 - 0{,}277 = 0{,}723.$$

Decisão Final: Aprovado ou Negado? Como $P(\text{Ruim} \mid E) = 0{,}723$ é maior que $P(\text{Bom} \mid E) = 0{,}277$, a decisão do modelo seria **Negar o empréstimo de Maria**.

Como a Hipótese de Independência Impacta o Resultado?

O modelo Naive Bayes assume que todas as variáveis (*Histórico com o banco*, *Estado Civil*, *Tempo de Emprego* e *Posse de Casa*) são independentes entre si. Na realidade, essas variáveis podem ter alguma correlação (por exemplo, pessoas casadas podem ter maior estabilidade financeira), mas o Naive Bayes ignora essas relações e trata cada fator separadamente.

Mesmo com essa suposição simplificadora, o modelo funciona bem em muitos problemas práticos, especialmente quando se tem um grande número de dados de treinamento. No entanto, se as variáveis forem altamente dependentes, o Naive Bayes pode ter um desempenho inferior a outros modelos mais sofisticados, como árvores de decisão ou redes neurais.

Conclusão

Com base nos cálculos, a probabilidade de Maria ter o empréstimo aprovado é de aproximadamente 27,7%, enquanto a probabilidade de negação é 72,3%. Como a classe com maior probabilidade é **"Ruim"**, o modelo Naive Bayes sugere que o empréstimo de Maria deve ser **Negado**.

4.3 Jogos Curiosos

4.3.1 Aplicação 07: O Problema de Monty Hall

O problema de Monty Hall é um paradoxo clássico da teoria das probabilidades e da estatística, baseado no programa de televisão americano *Let's Make a Deal* (Vamos fazer um trato), apresentado por Monty Hall. No Brasil, programas com a mesma ideia foram apresentados por Sergio Malandro (Porta dos desesperados) e por Silvio Santos (Porta da Esperança). O problema é um exemplo prático da aplicação do **Teorema de Bayes** e ilustra como a intuição humana pode falhar ao lidar com probabilidades condicionais.

Enunciado do Problema

No jogo, um participante enfrenta três portas fechadas. Atrás de uma delas há um prêmio valioso (um carro), e atrás das outras duas há prêmios indesejados (bodes).

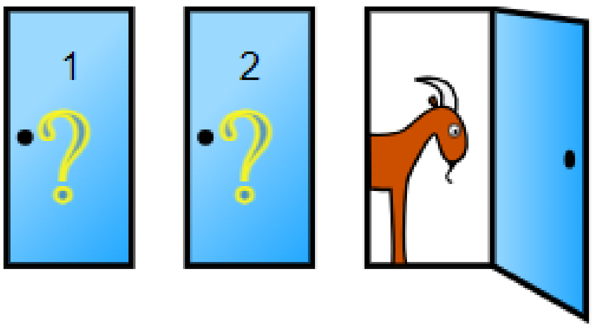

Figura 8 – Ilustração da ideia do problema
Fonte: pt.wikipedia.org/wiki/Problema de Monty Hall

O jogo segue as seguintes etapas:

1. O participante escolhe uma das três portas (sem abrir).
2. O apresentador, que sabe onde está o carro, abre uma das portas restantes, revelando um bode.
3. Agora, o participante deve tomar uma decisão:
   - **Manter** a porta inicialmente escolhida.

- **Trocar** para a outra porta que ainda está fechada.

A pergunta que surge é: **Qual estratégia maximiza a chance de ganhar o carro?**

Análise Probabilística

Inicialmente, a probabilidade de o jogador escolher a porta correta (onde está o carro) é:

$$P(\text{Carro na porta escolhida}) = \frac{1}{3}.$$

Consequentemente, a probabilidade de o carro estar em uma das outras duas portas é:

$$P(\text{Carro não está na porta escolhida}) = \frac{2}{3}.$$

Quando o apresentador abre uma porta com um bode, ele fornece uma informação valiosa: *a outra porta fechada agora tem toda a probabilidade que estava distribuída entre as duas portas restantes*. Isso significa que:

$$P(\text{Carro na outra porta} \mid \text{O apresentador abriu um bode}) = \frac{2}{3}.$$

Já a probabilidade de o carro estar na porta inicial permanece $1/3$.

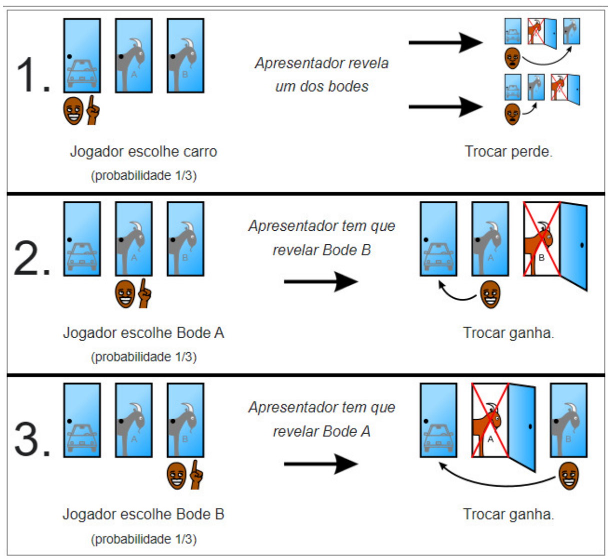

Figura 9 – Ilustração das possibilidades
Fonte: pt.wikipedia.org/wiki/Problema de Monty Hall

Demonstração pelo Teorema de Bayes

Para deixar claro o uso de Bayes, vamos focar num cenário concreto. Suponha que você escolha inicialmente a Porta A. Queremos comparar:

1. $P($ Carro atrás de A | Monty abriu B$)$

2. $P($ Carro atrás de C | Monty abriu B$)$

Isto é, dados que você escolheu A e que Monty abriu B (mostrando um bode), qual a probabilidade de o carro estar em A ou em C?

Passo 1: Definir os eventos

- Seja A o evento "o carro está atrás da Porta **A** ".

- Seja B o evento "o carro está atrás da Porta B ".

- Seja C o evento "o carro está atrás da Porta C".

- Seja M_B o evento "Monty abre a Porta B (com um bode)".

As probabilidades iniciais (antes de Monty abrir qualquer porta) são:

$$P(A) = P(B) = P(C) = \frac{1}{3}$$

Probabilidades condicionais de Monty abrir B

$$P(M_B \mid A), \quad P(M_B \mid B), \quad P(M_B \mid C).$$

- Se o carro estiver em A, Monty pode abrir B ou C com igual probabilidade (pois ambas contêm bode):
$$P(M_B \mid A) = \frac{1}{2}.$$
- Se o carro estiver em B, Monty não abrirá B (lá está o carro) nem A (sua escolha inicial). Então ele *obrigatoriamente* abriria C, o que implica:
$$P(M_B \mid B) = 0.$$
- Se o carro estiver em C, Monty obrigatoriamente abre B (pois A foi escolhida e C tem o carro):
$$P(M_B \mid C) = 1.$$

Aplicando a Fórmula de Bayes

Desejamos calcular $P(A \mid M_B)$. Pela Fórmula de Bayes:

$$P(A \mid M_B) = \frac{P(M_B \mid A)\, P(A)}{P(M_B \mid A)\, P(A) \; + \; P(M_B \mid B)\, P(B) \; + \; P(M_B \mid C)\, P(C)}.$$

Substituindo os valores:

$$P(A \mid M_B) = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{3}\right)}$$

$$= \frac{\frac{1}{6}}{\frac{1}{6} + 0 + \frac{1}{3}} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{2}{6}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.$$

Como a soma de todas as probabilidades deve ser 1, temos:

$$P(C \mid M_B) = 1 - P(A \mid M_B) = 1 - \frac{1}{3} = \frac{2}{3}.$$

Conclusão

- Manter a escolha inicial (Porta A): probabilidade de ganhar o carro $= \frac{1}{3}$.
- Trocar de porta (Porta C, no exemplo dado): probabilidade de ganhar o carro $= \frac{2}{3}$.

Portanto, a melhor estratégia para maximizar a chance de ganhar o carro é **sempre trocar** após Monty revelar um bode.

4.3.2 Aplicação 08: O Problema dos Três Prisioneiros

Problema proposto pela BBC News Brasil aos Leitores

Três prisioneiros estão presos e têm a chance de ganhar a liberdade se resolverem um desafio lógico.

O carcereiro traz **cinco chapéus**: **três brancos** e **dois vermelhos**. No escuro, ele coloca um chapéu na cabeça de cada prisioneiro, sem que eles saibam sua própria cor.

Depois disso, os prisioneiros são levados a um local iluminado. Dois deles podem ver os chapéus dos outros, mas não o seu próprio. O terceiro prisioneiro é cego e não pode ver nada.

O carcereiro então faz a seguinte pergunta:

1. O primeiro prisioneiro com visão é questionado:

   "Você sabe a cor do seu chapéu?"

   Ele responde: **"Não sei."**

2. O segundo prisioneiro, que também pode ver, recebe a mesma pergunta:

   "Você sabe a cor do seu chapéu?"

Ele também responde: **"Não sei."**

3. Finalmente, o prisioneiro cego é questionado:

"Você sabe a cor do seu chapéu?"

Após um momento de reflexão, ele responde com certeza: **"Sim, meu chapéu é branco!"**

Pergunta: Como o prisioneiro cego conseguiu deduzir corretamente a cor do seu chapéu?

Análise Probabilística

Temos 3 chapéus brancos e 2 vermelhos. A distribuição inicial de probabilidades para qualquer prisioneiro receber um chapéu é:

$$P(\text{chapéu branco}) = \frac{3}{5}, \quad P(\text{chapéu vermelho}) = \frac{2}{5}.$$

Etapa 1: O primeiro prisioneiro diz "Não Sei" Se o primeiro prisioneiro tivesse visto dois chapéus vermelhos, ele poderia concluir que seu próprio chapéu era branco, pois há apenas dois chapéus vermelhos disponíveis. Como ele disse **"Não sei"**, isso significa que ele não viu dois chapéus vermelhos. Ou seja, **pelo menos um dos outros dois tem um chapéu branco**.

Etapa 2: O segundo prisioneiro também diz "Não Sei" Agora, esse segundo prisioneiro tem mais informação: - Se ele tivesse visto um chapéu vermelho no cego, ele poderia concluir que o seu próprio chapéu era branco. - Mas ele também responde **"Não sei"**, o que significa que ele também não viu um chapéu vermelho no cego.

Etapa 3: O prisioneiro cego deduz sua cor O prisioneiro cego percebe que os outros dois prisioneiros não conseguiram determinar suas próprias cores. Se ele estivesse usando um **chapéu vermelho**, o segundo prisioneiro teria visto um vermelho e poderia deduzir sua própria cor. Como isso não aconteceu, ele percebe que **ele só pode estar usando um chapéu branco**.

$$P(\text{chapéu branco} \mid \text{ambos disseram "não sei}) = 1.$$

Solução com o Teorema de Bayes

Podemos formalizar essa dedução usando o Teorema de Bayes.

Seja: - H_B a hipótese de que o prisioneiro cego tem um **chapéu branco**. - H_V a hipótese de que o prisioneiro cego tem um **chapéu vermelho**. - D o evento de que os dois prisioneiros que enxergam disseram "Não Sei".

Queremos calcular:

$$P(H_B \mid D) = \frac{P(D \mid H_B)P(H_B)}{P(D)}$$

Sabemos que: - Inicialmente, $P(H_B) = \frac{3}{5}$ e $P(H_V) = \frac{2}{5}$. - Se o cego tem um chapéu **branco**, os outros dois prisioneiros têm 2 brancos e 1 vermelho. Nesse caso, os dois sempre responderão "Não Sei". - Se o cego tem um chapéu **vermelho**, há 1 branco e 2 vermelhos. Isso significa que o segundo prisioneiro teria visto um vermelho e poderia responder corretamente.

Ou seja:

$$P(D \mid H_B) = 1, \quad P(D \mid H_V) = 0.$$

Substituindo na fórmula:

$$P(H_B \mid D) = \frac{(1) \times \frac{3}{5}}{(1) \times \frac{3}{5} + (0) \times \frac{2}{5}} = \frac{3}{3} = 1.$$

Portanto, a certeza de que o cego tem um chapéu branco é 100%.

Conclusão

Esse problema é um excelente exemplo de como a probabilidade condicional e o Teorema de Bayes podem ser usados para resolver desafios lógicos. Assim como no **Problema de Monty Hall**, a solução vem de raciocinar sobre a informação que os outros jogadores forneceram indiretamente.

Ideia do Problema

A **ausência de uma resposta** pode ser uma pista tão valiosa quanto a presença de uma resposta. **Inferências baseadas em conhecimento parcial** são essenciais para a tomada de decisões sob incerteza.

4.3.3 Aplicação 08: O Problema da Caixa de Bertrand

O artigo "A Caixa de Bertrand", publicado na Revista de Ciência Elementar (dezembro de 2020) por Santos C. Dias, trata de um clássico problema de probabilidade condicional proposto pelo matemático Joseph Bertrand em 1889. O objetivo principal do artigo é ilustrar como a intuição probabilística humana pode levar a erros, e como a correta aplicação do Teorema de Bayes e da redução do espaço amostral pode solucionar problemas aparentemente paradoxais.

1. Contexto do Problema

Existem três caixas idênticas, fechadas. Sabe-se que uma das caixas contém duas moe- das de ouro, outra duas de prata e a terceira uma de prata e uma de ouro. Após a escolha aleatória de uma das caixas, é extraída uma moeda que se verifica ser de ouro. Desconhecendo-se qual era o conteúdo inicial da caixa, pretende-se saber qual a probabilidade de a outra moeda, dessa mesma caixa, ser também de ouro

As configurações das caixas são as seguintes:

Caixa A: Contém duas moedas de ouro.

Caixa B: Contém duas moedas de prata.

Caixa C: Contém uma moeda de ouro e uma de prata.

Um participante escolhe aleatoriamente uma das caixas e retira uma moeda. Ele observa que a moeda retirada é de ouro. A questão é:

Qual é a probabilidade de que a outra moeda da mesma caixa também seja de ouro?

De acordo com Santos (2020) "a resposta $\frac{1}{2}$ é a que se obtém, mais frequentemente, ao problema da “Caixa de Bertrand” e é equivocada. O equívoco, tal como destacou o próprio Bertrand, está em assumir que a probabilidade de a moeda que ficou na caixa, ser de ouro é igual à probabilidade de ser de prata. Essa conclusão equivocada tem por base uma primeira conclusão correta, de que, se a moeda extraída era de ouro, a caixa escolhida terá sido ou a que tem duas moedas de ouro ou a que tem uma moeda de ouro e outra de prata. Mas, como veremos mais em pormenor a seguir, o facto de ter saído uma moeda de ouro atribui diferentes probabilidades de escolha a cada uma destas duas caixas".

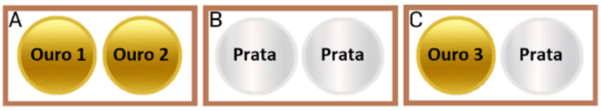

Figura 10 – Conteúdo das Caixas
Fonte: Santos (2020)

2. Solução Probabilística

A primeira abordagem para resolver esse problema é listar os possíveis cenários.

Se a moeda retirada era de ouro, então a Caixa B (com duas pratas) não pode ter sido escolhida.

Portanto, só existem duas possibilidades:

A Caixa A (duas de ouro) foi escolhida.

A Caixa C (ouro e prata) foi escolhida.

Agora, devemos calcular as probabilidades de cada uma dessas situações.

Passo 1: Probabilidade de Escolher Cada Caixa

Como a escolha da caixa foi aleatória, antes da extração da moeda, cada caixa tinha probabilidade igual de ser escolhida:

$$P(A) = P(B) = P(C) = \frac{1}{3}$$

Como já sabemos que a moeda retirada foi de ouro, eliminamos a possibilidade da Caixa B, e devemos atualizar as probabilidades para os casos restantes.

Passo 2: Probabilidade de Retirar Ouro em Cada Caixa

Sabemos que:

- Na Caixa A, ambas as moedas são de ouro, então a chance de retirar ouro ao acaso é 1 .

- Na Caixa C, há uma moeda de ouro e uma de prata, então a chance de retirar ouro ao acaso é $\frac{1}{2}$.

Logo, a probabilidade de retirar ouro em cada caixa

$$P(\text{ Ouro } \mid A) = 1, \quad P(\text{ Ouro } \mid C) = \frac{1}{2}$$

Como sabemos que a (primeira) moeda extraída foi de ouro, o acontecimento, $(Prata, Ouro_3)$ não está nas condições do problema, restando-nos três acontecimentos $(Ouro_1, Ouro_2)$, $(Ouro_2, Ouro_1)$ e $(Ouro3, Prata)$. Como, destes acontecimentos, apenas o primeiro e o segundo cumprem o que é pretendido - existir outra moeda de ouro na caixa de onde se retirou a primeira moeda de ouro, a probabilidade solicitada é de $\frac{2}{3}$ (Santos, 2020).

3. Solução pelo Teorema de Bayes

Agora aplicamos o Teorema de Bayes para calcular a probabilidade de que a caixa escolhida seja a Caixa A, dado que a moeda retirada foi de ouro:

$$P(A \mid \text{ Ouro }) = \frac{P(\text{ Ouro } \mid A)P(A)}{P(\text{ Ouro } \mid A)P(A) + P(\text{ Ouro } \mid C)P(C)}$$

Substituindo os valores:

$$
\begin{aligned}
P(A \mid \text{ Ouro }) &= \frac{(1) \times \frac{1}{3}}{(1) \times \frac{1}{3} + \left(\frac{1}{2} \times \frac{1}{3}\right)} \\
&= \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6}} = \frac{\frac{1}{3}}{\frac{2}{6} + \frac{1}{6}} = \frac{\frac{1}{3}}{\frac{3}{6}} = \frac{2}{3}.
\end{aligned}
$$

Portanto, a probabilidade de que a outra moeda da caixa escolhida também seja de ouro é $\frac{2}{3}$ (ou $66,7\%$).

4. Conclusão

O Problema da Caixa de Bertrand é um ótimo exemplo de como a intuição pode nos levar ao erro ao lidar com probabilidades condicionais. Muitas pessoas, sem fazer os cálculos corretamente, assumem que a probabilidade de a outra moeda ser de ouro seria $\frac{1}{2}$, mas a análise correta pelo Teorema de Bayes mostra que a resposta certa é $\frac{2}{3}$. Esse problema é similar ao Problema de Monty Hall, pois ambos demonstram que a revelação parcial de informações altera as probabilidades iniciais, um conceito fundamental em probabilidade condicional.

O Problema da Caixa de Bertrand mostra como a intuição pode nos levar a respostas incorretas. Muitas pessoas assumem que a outra moeda na caixa tem 50% de chance de ser ouro. Porém, a análise formal utilizando a redução do espaço amostral e o Teorema de Bayes mostra que a resposta correta é $\frac{2}{3}$

Esse problema está diretamente relacionado ao Problema de Monty Hall, pois ambos envolvem atualização de probabilidades após a revelação de uma informação parcial. Ele também tem variantes, como o jogo das três cartas e outras versões exploradas por Martin Gardner.

5 Considerações Finais

A probabilidade está presente em diversas áreas do cotidiano, mesmo que, muitas vezes, não a percebamos. Desde sua origem nos estudos de jogos de apostas, passando pelo desenvolvimento matemático nos séculos XVII e XVIII, até sua aplicação moderna na inteligência artificial e na estatística, ela impacta diretamente a forma como lidamos com a incerteza e tomamos decisões informadas, tornando-se uma ferramenta indispensável para a análise de fenômenos aleatórios.

Diante os conceitos básicos, o Teorema de Bayes nos mostra que a probabilidade de um evento acontecer pode ir sendo atualizada conforme novas informações que se vai sendo obtida. Por meio dele, é possível melhorar a precisão de previsões, aprimorar processos de tomada de decisão e aumentar a eficiência de diversas aplicações tecnológicas.

O algoritmo de Naive Bayes se destaca como uma ferramenta poderosa e eficiente para classificação em diversos domínios, desde a filtragem de spam até os diagnósticos médicos. Com base no Teorema de Bayes e na suposição de independência entre cada evento, reforçando a importância da estatística e da probabilidade na construção de sistemas inteligentes, a partir dos conhecimentos prévios.

A partir do avanço da inteligência artificial, reforça a necessidade de um maior domínio sobre os conceitos probabilísticos. Com a crescente quantidade de informações que estão sendo disponibilizadas e a necessidade de processá-las de forma eficiente, o raciocínio probabilístico se torna cada vez mais indispensável. Logo, ao investir no ensino acessível da probabilidade e do Teorema de Bayes não apenas beneficia o desenvolvimento acadêmico, mas também promove um pensamento analítico e crítico na alunos.

6 Referências

Dantas, C. A. B. Probabilidade: um curso introdutório. São Paulo: EDUSP, 2013. ISBN: 978-85-314-0399-6.

Bussab, W. O., Morettin, P. A. (2004). Estatística Básica (5ª ed.). São Paulo: Saraiva. ISBN: 85-02-03497-9.

JUNQUEIRA, Ana Lucia Nogueira. A Probabilidade que a História nos Conta. In: XIV Conferência Interamericana de Educação Matemática. Chiapas, México, 2015.

ALMEIDA, Alfredo Betâmio de. O Problema Epistemológico da Probabilidade e a contribuição de Karl Popper para o respectivo debate. Universidade Nova Lisboa, 2005. Disponível em: https://repositorio.ul.pt/bitstream/10451/36874/2/BETaMIODEALMEIDA. Acesso em: 08 mar. 2025

GOVERNO DO ESTADO DE MINAS GERAIS. Teste rápido Covid-19. Blog Coronavírus. Disponível em: https://coronavirus.saude.mg.gov.br/blog/68-teste-rapido-covid-19. Acesso em: 10 mar. 2025.

SELVIN, Steve. A Problem in Probability. The American Statistician, v. 29, n. 1, p. 67, 1975.

SANTOS, C.; DIAS, C. A caixa de Bertrand. Revista Ciência Elementar, v. 8, n. 4, p. 055, 2020. DOI: 10.24927/rce2020.055.

GARDNER, M. Ah, Apanhei-te!, Gradiva. 1993.

BBC BRASIL. Enigma matemático desafia milhares de internautas. BBC News Brasil. 16 abr. 2015. Disponível em: https://www.bbc.com/portuguese/noticias/2015/04/150416-enigma-bbc-pu. Acesso em: 06 mar. 2025.

LARREMORE, D. B. et al. Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. Science Advances, v. 7, n. 1, 2021. DOI: 10.1126/sciadv.abd5393.

MENDES, Iran Abreu; CHAQUIAM, Miguel. Um percurso pela história da probabilidade. ResearchGate, 2022. Disponível em: https://www.researchgate.net/publication/360985894-Um-percurso-pela-historia-da-probabilidade. Acesso em: 11 mar. 2025.

EVES, Howard. Introdução à história da matemática. Tradução: Hygino H. Domingues. 5. ed. Editora da Unicamp. Campinas, SP. 2011.

GELMAN, Andrew *et al. Bayesian Data Analysis.* 3rd ed. Chapman Hall/CRC, 2013.

SELVIN, Steve. *A Problem in Probability.* The American Statistician, v. 29, n. 1, p. 67, 1975.

VOS SAVANT, Marilyn. *Game Show Problem.* Parade Magazine, 9 de setembro de 1990.

FELLER, William. *An Introduction to Probability Theory and Its Applications.* Vol. 1. 3ª ed. Wiley, 1968.

LINK, Carlos Alberto. *Diagnóstico da gravidez ectópica usando o teorema de Bayes: estudo de corte retrospectivo.* 2022. 97 f. Dissertação (Mestrado em Ciências Cirúrgicas) - Universidade Federal do Rio Grande do Sul, 2022.

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA PARAÍBA
Campus Campina Grande - Código INEP: 25137409
R. Tranqüilino Coelho Lemos, 671, Dinamérica, CEP 58432-300, Campina Grande (PB)
CNPJ: 10.783.898/0003-37 - Telefone: (83) 2102.6200

Documento Digitalizado Ostensivo (Público)

Entrega de trabalho de conclusão de curso

Assunto:	Entrega de trabalho de conclusão de curso
Assinado por:	Sara Lopes
Tipo do Documento:	Requerimento
Situação:	Finalizado
Nível de Acesso:	Ostensivo (Público)
Tipo do Conferência:	Cópia Simples

Documento assinado eletronicamente por:

- **Sara Lopes Silva**, **DISCENTE (202111230007) DE LICENCIATURA EM MATEMÁTICA - CAMPINA GRANDE**, em 21/03/2025 13:41:24.

Este documento foi armazenado no SUAP em 21/03/2025. Para comprovar sua integridade, faça a leitura do QRCode ao lado ou acesse https://suap.ifpb.edu.br/verificar-documento-externo/ e forneça os dados abaixo:

Código Verificador: 1429355
Código de Autenticação: e0e4ee351e