UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO CIÊNCIA EXATA E DA TERRA
DEPARTAMENTO DE MATEMÁTICA MESTRADO PROFISSIONAL
EM MATEMÁICA EM REDE NACIONAL – PROFMAT

RENATO DOS SANTOS MAIA

O Teorema de Bayes: aplicações atuais

Orientador:

Prof. Dr. Marcelo Gomes Pereira

Natal/RN - 2021

RENATO DOS SANTOS MAIA

Teorema de Bayes: aplicações atuais

Dissertação apresentada ao Corpo Docente do Mestrado Profissional em Matemática em Rede Nacional - PROFMAT -CCET - UFRN, como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador:

Profº. Dr. Marcelo Gomes Pereira

Natal/RN - 2021

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Setorial Prof. Ronaldo Xavier de Arruda - CCET

Maia, Renato dos Santos.
O Teorema de Bayes: aplicações atuais / Renato dos Santos Maia. - 2021.
40f.: il.

Dissertação (Mestrado) - Universidade Federal do Rio Grande do Norte, Centro de Ciências Exatas e da Terra, Departamento de Matemática, Mestrado Profissional em Matemática em Rede Nacional - PROFMAT. Natal, 2021.
Orientador: Prof. Dr. Marcelo Gomes Pereira.

1. Matemática - Dissertação. 2. Teorema de Bayes - Dissertação. 3. Probabilidade condicional - Dissertação. 4. Covid-19 - Dissertação. I. Pereira, Marcelo Gomes. II. Título.

RN/UF/CCET CDU 51

Elaborado por Joseneide Ferreira Dantas - CRB-15/324

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
DEPARTAMENTO DE MATEMÁTICA
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL - PROFMAT

RENATO DOS SANTOS MAIA

Teorema de Bayes: aplicações atuais

Comissão Examinadora:

Prof°. Dr°. Marcelo Gomes Pereira (UFRN - Orientador)
Profª. Drª. Viviane Simioli Medeiros Campos (UFRN - Membro interno)
Prof°. Dr°. Ary Vasconcelos Medindo (UnB - Membro externo)

Natal/RN - 2021

Agradecimentos

Sou imensamente grato a todos que ajudaram de alguma maneira, para que esse momento fosse possível.

Dedicatória

Dedico este trabalho a minha esposa e meu filho.

Resumo

Vivemos em um mundo cada vez mais globalizado, onde cada vez mais recorremos aos dados estatísticos e probabilísticos para tomarmos decisões. E ao mesmo tempo estamos vivenciando situações completamente diferentes de tudo que havíamos visto até então, como a ***pandemia da COVID-19***. Neste trabalho, fazemos um estudo probabilístico relacionando o Teorema de Bayes à situações reais. Através de uma linguagem simples, e conceitualmente correta, podemos dar um toque mais atual aos conceitos de probabilidade estudados no Ensino Médio e quem sabe assim, atrair mais jovens à matemática.

PALAVRAS-CHAVE: Teorema de Bayes; Probabilidade condicional; covid-19.

Abstract

We live in an increasingly globalized world, where is increasingly common resort to statistical and probabilistic data to take decisions. At the same time, we are experiencing hypotheses that are different from anything we have seen so far, such as the ***covid-19 pandemic***. Thus, in this work, we will do a probabilistic study relating Bayes' Theorem to the covid, through Bayesian inference, and thus, we can give a more intriguing and current touch that can be used to attract more people to mathematics.

KEYWORDS : Bayes' theorem; bayesian inference; conditional probability; corona virus.

Sumário

Introdução

A Probabilidade é um tema que nos desperta interesse e curiosidade, quando nos deparamos com alguns jogos, muitos dos quais tem um fator ”sorte”. Pensamos em estratégias para vencer e mesmo que sem saber o que seja ou como trabalhar com a Probabilidade nós a utilizamos intuitivamente. Quando tomamos as mais diversas decisões no nosso cotidiano, ela continua lá. Até que alguém nos fala sobre esse termo, **Probabilidade**. Muitas vezes, a primeira vez que ouvimos é até de alguém que não sabe ao certo o que seja, mas mesmo assim ficamos maravilhados com as possibilidades de melhorar nossas decisões e o que podemos fazer com isso.

Uma pena que mais tarde, quando vemos o assunto na escola, muitas vezes nos deparamos com algo totalmente diferente do que esperávamos. Uma abordagem sem graça, que não nos motiva e entusiasma, até que acabamos de deixar aquilo de lado. É nessa hora que começa a apagar-se a figura da **curiosidade** dentro de nós.

Por outro lado, uma figura que nunca se apaga, pelo contrário, fica sempre mais forte a medida que deixamos de lado é a figura da **necessidade**. Todos nós temos necessidades, algumas em comum e outras não, mas todos temos. Como o próprio nome já diz: a necessidade é algo que não vivemos sem resolver, então nossa existência depende da nossa capacidade de resolvê-las (um baita motivador não?). Isso provavelmente te faz lembrar uma situação bem atual, a **Pandemia da Covid-19**.

É assim que vejo os dois pilares que movem nossas ações, principalmente estudos: **curiosidade** e **necessidade**. A necessidade é o que nos faz permanecer existindo e sobrevivendo enquanto a curiosidade é o que nos faz evoluir e criar. É por isso que neste trabalho será tratado estes 2 temas mencionados anteriormente, pois se faz necessário sabermos como lidar, como estão lidando com esta pandemia e também quero reacender (ou acender) a curiosidade sobre o estudo da Probabilidade àqueles que o lerem.

Desta maneira, o trabalho traz no primeiro capítulo, um paralelo entre a Probabilidade normalmente vista no Ensino Médio e a sua construção mais formal, são definidos e exemplificados conceitos como Dependência de eventos e Probabilidade condicional.

O Capítulo 2 trata sobre o Teorema de Bayes, com enunciado, versões alternativas, demonstração e exemplos em diferentes casos. Destaque ao Exemplo 2.0.2 que nos dá uma prévia do que vem a seguir no Capítulo 3.

Já o Capítulo 3 é onde estão as aplicações mais atuais do Teorema de Bayes, com um destaque à aplicação aos Testes da Covid, calculando a probabilidade do teste em questão ser falso-negativo. E o capítulo se encerra com uma outra aplicação relativa à Covid, desta vez relativa a medição de temperatura para se entrar em estabelecimentos.

Capítulo 1

Introdução à probabilidade

Quando nos deparamos com alguns jogos, pensamos em como podemos ganhar e quais são nossas possibilidades de resultados favoráveis do jogo. Por exemplo, se for uma moeda jogada ao acaso, teremos duas possibilidades: cara (c) e coroa (k). Porém, se jogarmos duas vezes a mesma moeda, nossas possibilidades já se alteram totalmente, pois teremos *cc, ck, kc* ou *kk*, ou seja, quatro possibilidades. A ideia se repete num dado cúbico comum, que em um lançamento teremos 6 possibilidades ($1, 2, 3, 4, 5$ ou 6). Mas se quisermos a soma de dois dados, teremos 11 possibilidades, são elas: $2, 3, 4, 5, 6, 7, 8, 9, 10, 11$ ou 12, daí já poderemos verificar que a soma ser igual a 2 ou 12 é bem mais difícil do que ser igual a sete, como veremos adiante nos exemplos do texto.

De fato, esse pensamento sobre possibilidades de vitória em um jogo não vem de hoje. No século XVII, os matemáticos franceses Blaise Pascal e Pierre de Fermat, atendendo a solicitação de alguns jogadores, começaram um estudo matemático desses jogos e com esses estudos, criaram uma nova teoria, que veio a ser chamada de teoria das probabilidades.

O objetivo deste capítulo é apresentar tópicos da teoria de probabilidades, visando expor conceitos de probabilidade a nível superior em uma linguagem simples e acessível ao ensino médio.

1.1 Probabilidade

Nesta seção definiremos nossos primeiros conceitos sobre probabilidade.

Cotidianamente nos deparamos com situações e acontecimentos que ao serem repetidos nas mesmas condições produzem resultados diferentes que não podem ser determinados previamente, por exemplo: Qual será o último número da placa do próximo carro que vai passar na rua? Pode ser $\{0, 1, \cdots, 9\}$. Isso nos leva a definição de *experimento aleatório*:

Definição 1.1.1. *Um experimento cujo o resultado depende* ***exclusivamente*** *do acaso é chamado de* ***experimento aleatório****. Ou seja, experimentos que apresentam resultados imprevisíveis quando repetidos, mesmo que as condições sejam semelhantes.*

O exemplo a seguir é o mais simples e, por isso, o mais utilizado para representar um experimento aleatório.

Exemplo 1.1.1. *O resultado do lançamento de uma moeda (o resultado pode ser cara ou coroa). Geralmente, não há como interferir no resultado do lançamento, nem prevê-lo.*

Formalmente:

Definição 1.1.2. *O conjunto de todos os resultados possíveis (****pontos-amostrais****) de um experimento aleatório é chamado de* ***espaço amostral*** *deste experimento. Normalmente, denotamos o espaço amostral por Ω.*

Para melhor entender a definição, vejamos os exemplos a seguir:

Exemplo 1.1.2. *No lançamento de um dado o espaço amostral é $\Omega = \{1, 2, 3, 4, 5, 6\}$.*

Exemplo 1.1.3. *Também podemos utilizar teoria dos conjuntos para facilitar a nossa vida. Vejamos quando o experimento é o lançamento de duas moedas (uma seguida da outra). Como os pontos-amostrais, isto é, os resultados possíveis, do lançamento de uma moeda são c ou k, terei que o espaço amostral Ω para o lançamento de 2 moedas é o produto cartesiano.*

Sendo A o conjunto dos resultados possíveis lançamento da primeira moeda e B o conjunto dos resultados possíveis do lançamento da segunda moeda:

$$A = \{c, k\} \text{ e } B = \{c, k\}$$

Assim, podemos escrever

$$\Omega = A \times B = \{(c,c),(c,k),(k,c),(k,k)\}.$$

Dentre todos os resultados possíveis (espaço amostral), podemos nos interessar por somente uma parte deles. Isso nos leva à definição de *"evento"* a seguir.

Definição 1.1.3. *Chamamos de* ***evento*** *um subconjunto de um espaço amostral.*

Exemplo 1.1.4. *Para o exemplo 1.1.2, vemos que* $\{1,2\}$, $\{2,3,4\}$, $\{5\}$ *são eventos deste espaço amostral.*

Definição 1.1.4. *Dizemos que os eventos* $A_1, A_2, \cdots, A_n$ *formam uma* ***partição do espaço amostral*** Ω*, se eles não tem interseção entre si e se sua união é igual ao espaço amostral. Isto é,*

$$A_i \cap A_j = \emptyset \text{ para } i \neq j \quad e \quad \bigcup_{i=1}^{n} A_i = \Omega.$$

Alguns tipos de eventos bem característicos:

Definição 1.1.5. *O* ***evento certo*** *é aquele evento que possui todos os elementos do espaço amostral. Por sua vez, o evento* ***impossível*** *é aquele evento igual ao conjunto vazio.*

Exemplo 1.1.5. *No jogar de um dado, o evento certo é o conjunto que representa obter resultado 1, 2, 3, 4, 5 ou 6, ou seja, todo o espaço amostral.*

Já se quisermos que ao jogar o dado apareça o número 7*, esse resultado representa o evento impossível, uma vez que 7 não está no espaço amostral.*

Definição 1.1.6. *Se* A *é um evento de um espaço amostral* Ω*, o evento* ***complementar*** *de* A*, indicado por* A^c *é* $A^c = \Omega - A$*.*

Exemplo 1.1.6. *Ao jogar um dado, o espaço amostral do experimento será* $\{1,2,3,4,5,6\}$*. O evento "resultado par" é* $P = \{2,4,6\}$ *e o evento "resultado ímpar" é* $I = \{1,3,5\}$*.*

Note que $I = \{1,3,5\} = \{1,2,3,4,5,6\} - \{2,4,6\} = \{2,4,6\}^c = P^c$*, ou seja,* I *é o evento complementar a* P*.*

Neste ponto, vale lembrar que estabelecemos uma conexão entre os resultados possíveis de um experimento aleatório e subconjuntos, que chamamos de eventos. Por vezes, usaremos a expressão *eventos aleatórios* levando em conta essa relação, porém pode ser lido apenas como eventos, se preferir. Na observação dos resultados de um experimento, surge a necessidade de quantificar, ou de estabelecer uma medida para, o grau de dificuldade de obtenção dos eventos. Essa medida é o que chamamos de probabilidade e que será a definida formalmente a seguir.

Mas antes, a definição de um caso particular, que é a definição usual nos livros do Ensino Médio, como podemos observar em [11]. Um espaço amostral é chamado **equiprovável** quando todos seus resultados possíveis (pontos amostrais) têm a mesma chance (grau de dificuldade) de ocorrer. É o caso de lançamentos de dados, ou de moedas, não viciados, escolha de bolas numeradas de tamanho e peso idênticos etc. Muitas vezes, é a primeira aproximação que se faz em um experimento qualquer. Quando se ouve falar em um concurso público com certa quantidade de candidatos, por exemplo, a primeira reação é achar que todos têm a mesma chance de conseguir aprovação, ou seja, o espaço amostral para aprovação do concurso é equiprovável. Essa ideia muda se levarmos em consideração os diferentes níveis de preparação de cada candidato, pois alguns se preparam mais do que outros. Esses têm maiores chances. Outro exemplo, mais simples, de espaço amostral que pode ser considerado não equiprovável é do seguinte experimento: escolher entre tomar sorvete ou fazer caminhada em um dia de sol.

Definição 1.1.7. *Dado um experimento aleatório, sendo Ω o seu espaço amostral* ***finito*** *e equiprovável. Definimos a probabilidade de um evento $A \subset \Omega$ ao número real $P(A)$ tal que:*

$$P(A) = \frac{card(A)}{card(\Omega)},$$

onde $card(A), card(\Omega)$ significam a quantidade de elementos do conjunto A e Ω, respectivamente.

Exemplo 1.1.7. *Qual a probabilidade de obtermos um número par no lançamento de um dado honesto?*

Os números pares possíveis em um dado são $2, 4$ *e* 6*, ou seja, estamos interessados no evento* $A = \{2, 4, 6\} \Rightarrow card(A) = 3$ *e sabemos que nosso espaço amostral é* $\Omega = \{1, 2, 3, 4, 5, 6\} \Rightarrow card(\Omega) = 6$*. Assim:*

$$P(A) = \frac{card(A)}{card(\Omega)} = \frac{3}{6} = \frac{1}{2}.$$

Notemos que as probabilidades sempre resultarão em um número dentro do intervalo $[0, 1]$, pois, A é um subconjunto de Ω, tendo assim, no mínimo nenhum elemento e, no máximo, a quantidade de elementos de Ω.

A curiosidade de vermos os resultados favoráveis ou não favoráveis é um dos motivos que geram um grande interesse na disciplina de probabilidade como vimos na introdução. Note que para calcularmos as probabilidades de sucesso de um experimento aleatório acontecer, precisamos conhecer todos os seus resultados possíveis (*espaço amostral*). Por exemplo, no caso do lançamento de uma moeda temos $\Omega = \{c, k\}$, já no caso de jogarmos 2 moedas $\Omega = \{cc, ck, kc, kk\}$ e assim sucessivamente.

Em seguida, estendemos a definição de probabilidade para torná-la mais geral, onde o espaço amostral Ω não é necessariamente equiprovável e finito.

Definição 1.1.8. *Para um espaço amostral* Ω*, dizemos que um conjunto* $\mathcal{F}$ *formado por subconjuntos de* Ω *que satisfazem*

1. $\Omega \in \mathcal{F}$.
2. *Se* $A \in \mathcal{F}$*, então* $A^c \in \mathcal{F}$.
3. *Se* $A_1, A_2, \cdots \in \mathcal{F}$*, então* $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

é uma ***classe de eventos aleatórios****. Os elementos desta classe chamamos de* ***eventos aleatórios****.*

Exemplo 1.1.8. *No caso do lançamento de uma moeda com* $\Omega = \{c, k\}$ *temos*

$$\mathcal{F} = \{\{c\}, \{k\}, \{c, k\}, \emptyset\}$$

é uma classe de eventos aleatórios.

Definição 1.1.9. *Uma probabilidade é uma função $P : \mathcal{F} \longrightarrow [0,1]$ que satisfaz:*

1. $0 \leq P(A) \leq 1 \quad \forall A \in \mathcal{F}$.
2. $P(\Omega) = 1$.
3. *Se $A_1, A_2, \cdots \in \mathcal{F}$ e $A_i \cap A_j = \emptyset$ sempre que $i \neq j$, então*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Vejamos como a Definição 1.1.7 é um caso particular da Definição 1.1.9 em que o espaço amostral é **finito** e **equiprovável**:

Como Ω é finito, denotaremos $\Omega = \{a_1, a_2, \cdots, a_n\}$, ou seja, $\mathrm{card}(\Omega) = n$ e $\mathcal{F}$ o conjunto das partes de Ω, que claramente obedece a definição de classe de eventos aleatórios 1.1.8. Supondo que os eventos são equiprováveis temos

$$\begin{aligned}
1 &= P(\Omega) \quad (\text{ item } 2) \\
&= P\left(\{a_1, a_2, \cdots, a_n\}\right) \\
&= P\left(\{a_1\}\right) + \cdots + P\left(\{a_n\}\right) \qquad (\text{item } 3) \\
&= n \cdot P\left(\{a_i\}\right), \quad (\text{eventos equiprováveis})
\end{aligned}$$

$$\Rightarrow P\left(\{a_i\}\right) = \frac{1}{n},$$

onde a_i é qualquer um dos elementos de Ω. Logo, se tomarmos $A \in \mathcal{F}$ temos que $A = \{a_{i_1}, a_{i_2}, \cdots, a_{i_k}\}$ para alguns $i_1, \cdots, i_k \in \{1, \cdots, n\}$, logo, $\mathrm{card}(A) = k$.

Assim,

$$A = \{a_{i_1}\} \cup \{a_{i_2}\} \cup \cdots \cup \{a_{i_k}\}$$
$$\Rightarrow P(A) = \textstyle\sum_{j=1}^{k} P\left(\left\{a_{i_j}\right\}\right) = k \cdot P\left(a_{i_1}\right) = k \cdot \dfrac{1}{n} = \frac{k}{n} = \dfrac{\mathrm{card}(A)}{\mathrm{card}(\Omega)}.$$

Vejamos então como trabalhamos com probabilidade utilizando a Definição 1.1.7.

Para calcular a probabilidade de um evento ocorrer, basta então dividirmos o número de ocorrências deste tal evento dentro do espaço amostral finito e equiprovável pela quantidade total do espaço amostral, vejamos os exemplos.

Exemplo 1.1.9. *Qual a probabilidade de ao jogar uma moeda (não viciada) o resultado ser c (cara)?*

Espaço amostral: $\Omega = \{k, c\} \Rightarrow card(\Omega) = 2$.

Probabilidade de ser cara: $P(\{c\}) = \dfrac{card(\{c\})}{card(\Omega)} = \dfrac{1}{2}$.

Exemplo 1.1.10. *Qual a probabilidade de ao jogarmos 2 dados a soma dos resultados dos lançamentos ser 7?*

Podemos utilizar o produto cartesiano para nosso espaço amostral, como já vimos anteriormente. Considerando

$$A = \{1, 2, 3, 4, 5, 6\} \text{ e } B = \{1, 2, 3, 4, 5, 6\}$$

temos que

$$\Omega = A \times B = \{(1, 1), (1, 2), (1, 3), \ldots, (6, 5), (6, 6)\}$$

e assim $card(\Omega) = card(A \times B) = card(A) \cdot card(B) = 6 \cdot 6 = 36$.

Os eventos em que a soma é 7 são $A = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$, *logo,* $card(A) = 6$.

Portanto,

$$P(A) = \frac{6}{36} = \frac{1}{6}.$$

1.2 Probabilidade condicional

Como temos o interesse em estudar o Teorema de Bayes, precisamos entender sobre probabilidade condicional. A probabilidade condicional refere-se à probabilidade de um evento A, sabendo que ocorreu um outro evento B. Para isso, usamos a notação $P(A|B)$. Lê-se ”P de A dado B”, ou ainda, ”P de A dado que B ocorreu”.

Definição 1.2.1. *Sejam A, B eventos de um espaço amostral Ω tais que $P(B) > 0$*

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

Caso $P(B) = 0$, B não pode ocorrer, o que torna a expressão acima indefinida e nesse caso, dizemos que $P(A|B) = 0$. Um exemplo do uso da Probabilidade Condicional: qual a probabilidade de, em jogando um dado, a face voltada para cima seja 2, sabendo que essa face é par? A informação prévia da ocorrência do evento B pode ser interpretada como espaço amostral para o cálculo da probabilidade.

Originalmente teríamos:

$$\Omega = \{1, 2, \ldots, 6\}$$

$$P(\{2\}) = \frac{1}{6}$$

Como sabemos que o dado foi par, teremos:

$$\Omega = \{2, 4, 6\}$$

$$A = \{2\}, \quad B = \{2, 4, 6\}$$

Logo, $P(A) = \dfrac{card(A)}{card(\Omega)} = \dfrac{1}{3}$.

Resultado esse que equivale a definição 1.2.1, pois:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{2\})}{P(\{2, 4, 6\})} = \frac{1/6}{3/6} = \frac{1}{3}.$$

Exemplo 1.2.1. *Numa loja de brinquedos temos dois tipos de brinquedos, bolas e dados, em duas cores, azuis e vermelhas, distribuídas da seguinte maneira:*

- 8 *bolas, sendo* 5 *vermelhas e* 3 *azuis;*
- 6 *dados, sendo* 4 *vermelhos e* 2 *azuis.*

Queremos saber, se escolhendo aleatoriamente, qual a probabilidade do brinquedo ser vermelho, entretanto sabendo que o brinquedo escolhido foi uma bola.

Vamos considerar por V *o evento brinquedo vermelho e* B *o evento bola. Assim, estamos interessados em calcular* $P(V|B)$*.*

$$P(V|B) = \frac{P(V \cap B)}{P(B)}$$

$$card(V) = 9, \quad card(b) = 8, \quad card(\Omega) = 14, \quad card(V \cap B) = 5$$

$$P(B) = \frac{card(B)}{card(\Omega)} = \frac{8}{14}$$

$$P(V \cap B) = \frac{card(V \cap B)}{card(\Omega)} = \frac{5}{14}$$

$$P(V|B) = \frac{\frac{5}{14}}{\frac{8}{14}} = \frac{5}{14} \cdot \frac{14}{8} = \frac{5}{8}.$$

Probabilidade de interseções:

O cálculo da probabilidade de eventos que acontecem ao mesmo tempo, os chamados eventos simultâneos.

A fórmula para o cálculo dessa probabilidade decorre da fórmula da probabilidade condicional. Assim, teremos:

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

$$\Rightarrow P(A \cap B) = P(A|B) \cdot P(B)$$

Em palavras: a probabilidade de ocorrer $A \cap B$ é igual ao produto da probabilidade de um deles pela probabilidade do outro dado o primeiro.

Indutivamente:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Ou seja, para calcularmos uma porção de probabilidades, sabendo uma a uma o resultado.

Teorema 1. ***[Teorema do produto]*** *Sejam $A_1, \cdots, A_n \in \Omega$ com $P\left(\bigcap_{i=1}^{n} A_i\right) > 0$, então*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Demonstração. Faremos a demonstração utilizando o princípio da indução finita.

- **Caso $n = 2$:** Da definição de probabilidade condicional temos que $P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1)$, provando o caso $n = 2$.

- Passo indutivo:

  Hipótese da indução: $P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1})$

  Tese da indução: $P(A_1 \cap A_2 \cap \cdots \cap A_k \cap A_{k+1}) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{k-1})P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{k-1} \cap A_k)$

  De fato, basta aplicar o caso $n = 2$ e em seguida a hipótese de indução:

  $$P((A_1 \cap A_2 \cap \cdots \cap A_k) \cap A_{k+1}) = \underbrace{P(A_1 \cap A_2 \cap \cdots \cap A_k)}_{\text{H.I.}} P(A_k|A_1 \cap A_2 \cap \cdots \cap A_k)$$

  $$= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_k|A_1 \cap \cdots \cap A_{k-1})P(A_{k+1}|A_1 \cap \cdots \cap A_k).$$

$\square$

Se os eventos A e B forem independentes, ou seja, se o fato de ocorrer o evento B não alterar a probabilidade de ocorrer o evento A, a fórmula para o cálculo da probabilidade condicional será

$$P(A \cap B) = P(A) \cdot P(B).$$

Escrevendo isso como uma definição:

Definição 1.2.2. *Dois eventos A e B são ditos independentes quando $P(A \cap B) = P(A) \cdot P(B)$..*

Vamos fazer alguns exemplos para explorar o uso da fórmula e a maneira correta de interpretar os problemas relacionados à probabilidade de eventos simultâneos.

Exemplo 1.2.2. *Em dois lançamentos sucessivos de um mesmo dado, qual a probabilidade de ocorrer um número maior que 3 e o número 2?*

Perceba que a ocorrência de um evento não influencia a probabilidade de outro ocorrer, portanto são dois eventos independentes. Vamos distinguir os dois eventos:

$A = \{4, 5, 6\}$
$B = \{2\}$
Vamos calcular a probabilidade de ocorrência de cada um dos eventos.
Observe que no lançamento de um dado temos 6 valores possíveis. Assim:
$P(A) = \frac{3}{6} = \frac{1}{2}$ *e* $P(B) = \frac{1}{6}$
Dessa forma,

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}.$$

Exemplo 1.2.3. *Numa urna há 30 bolinhas numeradas de 1 a 30. Serão retiradas dessa urna duas bolinhas, ao acaso, uma após a outra, sem reposição. Qual a probabilidade de sair um múltiplo de 10 na primeira e um número impar na segunda?*

Solução. *O fato de a retirada das bolinhas ocorrer sem reposição, implica que a ocorrência do primeiro evento interfere na probabilidade do segundo ocorrer. Portanto esses eventos não são independentes.*

$\Omega = \{(a, b) \,|\, a, b \in \{1, 2, \cdots, 30\}\}$
Vamos determinar cada um dos eventos:
A*: sair um múltiplo de* 10 *na primeira retirada.*
B*: sair um número impar na segunda retirada.*

A probabilidade de ocorrer os dois eventos sucessivamente será dada por $P(A \cap B)$.

Para o cálculo de $P(B|A)$ *é preciso notar que tendo ocorrido* A*, não teremos mais* 30 *bolinhas na urna, pois uma foi retirada e não houve reposição, restando 29 bolinhas na urna. Assim,*

$$P(A \cap B) = P(B|A) \cdot P(A) = \frac{15}{29} \cdot \frac{1}{10} = \frac{3}{58}.$$

Exemplo 1.2.4. *Se tivermos* 6 *bolas, sendo* 4 *vermelhas e* 2 *azuis. Qual a probabilidade de retirarmos a primeira vermelha, a segunda azul e a terceira vermelha?*

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)$$

$$P(A_1) = \frac{4}{6}, \quad P(A_2|A_1) = \frac{2}{5}$$

$$P(A_3|A_1 \cap A_2) = \frac{3}{4}$$

Logo:

$$P(A_1 \cap A_2 \cap A_3) = \frac{4}{6} \cdot \frac{2}{5} \cdot \frac{3}{4} = \frac{1}{5}.$$

Probabilidade total:

Sejam $\{A_1, A_2, \ldots, A_n\}$ uma partição da Ω e B um evento de Ω.

Segue que:

$$\begin{array}{rcccccccc} P(B) & = & P(A_1 \cap B) & + & P(A_2 \cap B) & + & \ldots & + & P(A_n \cap B) \\ & = & P(A_1)P(B|A_1) & + & P(A_2)P(B|A_2) & + & \ldots & + & P(A_n)P(B|A_n). \end{array}$$

Teorema 2. ***[Teorema da probabilidade total]*** *Seja $\{A_1, \cdots, A_n\}$ uma partição de Ω com $P(A_i) > 0 \quad \forall i \in \{1, \cdots, n\}$. Então, para todo evento B*

$$P(B) = \sum_{i-1}^{n} P(A_i)P(B \mid A_i).$$

Demonstração. Pelo teorema do produto temos que $P(A_i)P(B|A_i) = P(A \cap B_i)$ para todo i de 1 até n. Como $\{A_1, \cdots, A_n\}$ é uma partição de Ω, segue $(B \cap A_i) \cap (B \cap A_j) = \emptyset$ para $i \neq j$. Assim,

$$\begin{array}{rcl} \sum_{i=1}^{n} P(A_i)P(B|A_i) & = & \sum_{i=1}^{n} P(B \cap A_i) \\ & = & P\left(\bigcup_{i=1}^{n}(B \cap A_i)\right) \\ & = & P\left(B \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) \\ & = & P(B). \end{array}$$

$\square$

A seguir, um exemplo de aplicação.

Exemplo 1.2.5. *Temos* 3 *lojas que vendem bolas, essas bolas são vermelhas ou azuis, a primeira loja tem* 8 *bolas, sendo* 5 *vermelhas e* 3 *azuis, a segunda loja tem* 10 *bolas, sendo* 6 *vermelhas e* 4 *azuis e a terceira loja tem* 12 *bolas, sendo* 2 *vermelhas e* 10 *azuis. Uma pessoa vai a uma das lojas comprar ao acaso uma bola. Sabendo que a pessoa gosta mais da terceira loja e que tem* 50% *de probabilidade de comprar na terceira loja, tem* 30% *de probabilidade de comprar na segunda e* 20% *de comprar na primeira, qual a probabilidade dela comprar uma bola azul?*

Sejam $A_1 = \{Loja1\}$, $A_2 = \{Loja2\}$, $A_3 = \{Loja3\}$ *e* $B = \{Azul\}$.

$$P(A_1) = \frac{20}{100}, \quad P(A_2) = \frac{30}{100}, \quad P(A_3) = \frac{50}{100}$$

$$P(B|A_1) = \frac{3}{8}, \quad P(B|A_2) = \frac{4}{10}, \quad P(B|A_3) = \frac{10}{12}$$

$$\begin{aligned} P(B) &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3) \\ &= \frac{20}{100} \cdot \frac{3}{8} + \frac{30}{100} \cdot \frac{4}{10} + \frac{50}{100} \cdot \frac{10}{12} \\ &= \frac{3}{40} + \frac{6}{50} + \frac{5}{12} = \frac{45 + 72 + 250}{600} = \frac{367}{600} \end{aligned}$$

$$\Rightarrow P(B) = \frac{367}{600} = \frac{\frac{367}{6}}{100} = 61,17\%.$$

Portanto, a probabilidade dessa pessoa comprar uma bola azul é de $61,17\%$.

1.2.1 Dependência e independência entre eventos

Alterar ou não a ocorrência de determinada situação em função de outra, permite que, a partir da relação entre dois eventos que podem ou não influenciar um ao outro, definamos a chamada ***relação de independência ou dependência entre eventos***, como já vimos anteriormente na definição 1.2.2.

Podemos também ver através da **probabilidade condicional quando dois eventos são ou não independentes**:

Como $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$, se A e B forem independentes teremos que $P(A \cap B) = P(A) \cdot P(B)$, logo,

$$P(A|B) = P(A)$$

Exemplo 1.2.6. *No lançamento de dois dados, qual a probabilidade de que tenhamos um número inferior a 3 no* 1° *lançamento e um número primo no* 2° *lançamento?*

Considerando como evento A, *o lançamento do primeiro dado, e como evento* B, *o lançamento do segundo dado, temos:*

$$\begin{cases} \Omega = \{(a,b) \,|\, a,b \in \{1,2,\cdots,6\}\} & \Rightarrow \ \#\Omega = 36 \\ A = \{(1,b),(2,b) \,|\, b \in \{1,2,\cdots,6\}\} & \Rightarrow \ \#A = 12 \\ B = \{(a,2),(a,3),(a,5) \,|\, a \in \{1,2,\cdots,6\}\} & \Rightarrow \ \#B = 18 \end{cases}$$

$$\Rightarrow P(A \cap B) = P(A).P(B) = \frac{12}{36} \cdot \frac{18}{36} = \frac{216}{1296} = \frac{1}{6}.$$

Exemplo 1.2.7. *Lançando-se simultaneamente um dado e uma moeda, determine a probabilidade de se obter* 3 *ou* 8 *no dado e cara na moeda.*

Separando os experimentos $\Omega_{dado} = \{1,2,\cdots,6\}$ *e* $\Omega_{moeda} = \{c,k\}$ *chegamos a* $\Omega = \Omega_{dado} \times \Omega_{moeda}$.

Considerando os eventos

$A = \{(1,c),(3,c),(1,k),(3,k)\}$ *e* $B = \{(1,c),(2,c),(3,c),(4,c),(5,c),(6,c)\}$.

Como $card(\Omega) = 12$, $card(A) = 4$ *e* $card(B) = 6$*:*

$P(A) = \dfrac{4}{12} = \dfrac{1}{3}$ *e* $P(B) = \dfrac{6}{12} = \dfrac{1}{2}$.

Logo,

$$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}.$$

Então teremos para moeda ser cara e o dado ser 1 ou 3, a probabilidade é de $\dfrac{1}{6}$.

Uma vez que entendemos as componentes Probabilidade, Probabilidade Condicional e Eventos Independentes, podemos nos dedicar ao Teorema de Bayes, ferramenta central a ser usada na dissertação, que será apresentado no capítulo seguinte.

Capítulo 2

Teorema de Bayes

O matemático Thomas Bayes foi um reverendo presbiteriano que viveu no século 18, entre os anos de 1702 e 1761, na Inglaterra. Na matemática, Bayes dedicou sua vida ao estudo das causas e efeitos da probabilidade, porém, jamais chegou a tornar público durante sua vida, nenhum resultado de seus estudos na área de probabilidade. A única publicação de Bayes na área da matemática se deu em 1936 com o livro chamado *"The doctrine of fluxion"* que tratava sobre a derivada de uma função contínua.

O seu amigo e filósofo Richard Price (1732-1791), apresentou a Real Sociedade Inglesa, uma entidade científica da época, em 1763, dois anos após a morte de Bayes, um artigo encontrado entre os papéis de Bayes, com o nome de *"An essay towards solving a problem in doctrine of chances"* (Ensaio buscando resolver um problema na doutrina das probabilidades). Este artigo tinha a demonstração do Teorema de Bayes.

0 Teorema de Bayes é uma das principais ferramentas que temos para *"refinar"* uma previsão probabilística (probabilidade a priori) quando novos eventos, em geral não independentes, acontecem. A probabilidade a posteriori é justamente o resultado desse *"refinamento"* frente às novas informações. E isso é útil, como por exemplo, perante uma nova doença como a covid-19 e com todas as incertezas que traz, afinal os testes serológicos (testes que procuram a presença de anticorpos contra o vírus) são imperfeitos, já que a precisão desta análise requer muitos fatores, incluindo a prevalência e raridade da doença. Por exemplo, no primeiro teste de anticorpos ao SARS-CoV-2 que foi aprovado pelo FDA (Food and Drug Administration) parecia estar errado tantas vezes quanto estava certo. Mas ao aplicar-se o teorema de Bayes,

tornou-se possível avaliar o que realmente se deseja saber: a probabilidade de o resultado do teste estar correto.

Tudo isso ajuda a explicar que muitas informações orientam o nosso comportamento no contexto da Covid-19 são probabilísticas. Por exemplo, sendo algumas estimativas, se formos infetados pelo coronavírus, temos um por cento de hipóteses de morrer. Mas a realidade é que a probabilidade de cada um tem uma enorme variação, dependendo da idade e de outros fatores.

O teorema de Bayes é um corolário da Lei da probabilidade total (Teorema 2), expresso matematicamente na forma da equação

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{2.1}$$

em que

- A e B são eventos e $P(B) \neq 0$;
- $P(A)$ e $P(B)$ são probabilidades a priori de A e B;
- $P(A|B)$ é a probabilidade a posteriori (probabilidade condicionada) de A condicional a B.

Formalmente temos:

Teorema 3. ***[Teorema de Bayes]*** *Seja* $\{A_1, \cdots, A_n\}$ *uma partição de* Ω *com* $P(A_i) > 0$ *para todo* i *entre* 1 *e* n*. Então, para todo evento* B*, com* $P(B) > 0$ *e para todo* j *entre* 1 *e* n*, tem-se que*

$$P(A_j \mid B) = \frac{P(B \mid A_j)P(A_j)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}. \tag{2.2}$$

Demonstração. Dividiremos a demonstração em dois casos: $P(A_j \cap B) = 0$ e $P(A_j \cap B) \neq 0$.

Primeiramente provaremos o caso $P(A_j \cap B) = 0$. Temos que

$$0 = P(A_j \cap B) = P(A_j \mid B)P(B)$$

e como $P(B) \neq 0$ segue que $P(A_j \mid B) = 0$. Como $A_j \cap B = B \cap A_j$, similarmente obtemos que $P(B \mid A_j) = 0$, uma vez que $P(A_i) > 0$ para todo i. Provando assim o teorema no caso $P(A_j \cap B) = 0$, uma vez que ambos os lados da igualdade são 0.

Agora, caso $P(A_j \cap B) \neq 0$. Como $P(A_j \cap B) \neq 0$ podemos utilizar o teorema do produto: $P(B \mid A_j)P(A_j) = P(A_j \cap B)$ e pelo teorema da probabilidade total temos $\sum_{i=1}^{n} P(B \mid A_i)P(A_i) = P(B)$, logo,

$$\frac{P(B \mid A_j)P(A_j)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)} = \frac{P(A_j \cap B)}{P(B)} = P(A_j \mid B).$$

$\square$

Mas talvez você esteja se perguntando: no começo foi falado que o teorema de Bayes era

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

e depois falou que era

$$P(A_j \mid B) = \frac{P(B \mid A_j)P(A_j)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)}.$$

Eles dizem a mesma coisa?

Sim! Basta considerar a partição $\{A, A^c\}$ de Ω e utilizar o teorema da probabilidade total: $P(B|A)P(A) + P(B|A^c)P(A^c) = P(B)$.

Devido a grande utilização dessa forma utilizando o complementar, diremos o seguinte:

Forma alternativa

Quando encontramos duas hipóteses concorrentes, isto é, hipótese A e hipótese concorrente A^c, usamos a seguinte fórmula:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^C)P(A^C)}$$

e dizemos que:

- $P(A^C)$ é a probabilidade correspondente do grau de crença inicial contra A. $P(A^C) = 1 - P(A)$
- $P(B|A^C)$ é a probabilidade condicional de B, dado a proposição A é falsa.

Exemplo 2.0.1. *Uma companhia multinacional tem três fábricas que produzem o mesmo tipo de produto. A fábrica I é responsável por 30% do total produzido, a fábrica II produz 45% do total, e o restante vem da fábrica III. Cada uma das fábricas, no entanto, produz uma proporção de produtos que não atendem aos padrões estabelecidos pelas normas internacionais. Tais produtos são considerados "defeituosos" e correspondem a 1%, 2% e $1,5\%$, respectivamente, dos totais produzidos por fábrica.*

No centro de distribuição, é feito o controle de qualidade da produção combinada das fábricas.

1. *Qual a probabilidade de encontrar um produto defeituoso durante a inspeção de qualidade?*

   Seja o evento $A = \{$Produto defeituoso$\}$ e $F_i = \{$Produto da fábrica i$\}$. Sabemos que $P(F_1) = 0,3$, $P(F_2) = 0,45$ e $P(F_3) = 0,25$. Além disso, sabemos que $P(A|F_1) = 0,01$, $P(A|F_2) = 0,02$ e $P(A|F_3) = 0,015$.

   Então, utilizando a Lei de probabilidade total,

   $$P(A) = P(A|F_1) \cdot P(F_1) + P(A|F_2) \cdot P(F_2) + P(A|F_3) \cdot P(F_3)$$

   $$\Rightarrow P(A) = 0,03 \cdot 0,01 + 0,45 \cdot 0,02 + 0,25 \cdot 0,015 = 0,01575.$$

2. *Se durante a inspeção encontrarmos um produto defeituoso, qual é a probabilidade que ele tenha sido produzido na fábrica II?*

   Aqui usaremos de fato o Teorema de Bayes:

   $$\begin{aligned} P(F_2|A) &= \frac{P(A|F_2) \cdot P(F_2)}{P(A)} \\ &= \frac{0,02 \cdot 0,45}{0,01575} \\ &= 0,5714 \end{aligned}$$

Exemplo 2.0.2. *Uma aplicação bem prática e atual do Teorema de Bayes é a solução do problema de interpretação do resultado de um teste de positividade sobre alguma doença.*

Antes de começar esse exemplo, você precisa ter conhecimento sobre uma coisa: ***testes não são 100% precisos!***

Imaginemos que o teste de mamografia se comporte da seguinte forma:

1. *1% das mulheres têm câncer de mama.*
2. *80% das mamografias detectam o câncer quando ele existe.*
3. *9,6% das mamografias detectam o câncer quando ele não existe.*

Dadas todas essas informações, imagine que você se submeteu ao teste de mamografia e esse teste apresentou um resultado positivo. Quais são as chances de realmente se ter câncer, dado que o teste deu ***positivo****?*

Analisemos os dados:

1. *Do* ***primeiro item****, se 1% das mulheres têm câncer de mama, então 99% não tem.*
2. *Do* ***segundo item****, se 80% das mamografias detectam o câncer quando ele existe, temos que 20% falha.*
3. *E do* ***terceiro item****, se 9,6% das mamografias detectam o câncer quando ele não existe, temos que 90,4% retornam corretamente um resultado negativo.*

Definamos o eventos $A = \{\text{tem câncer}\}$ *e* $B = \{\text{o teste deu positivo}\}$*, logo,* $P(A)$ *é a probabilidade de ter câncer e* $P(B)$ *a probabilidade do teste ser positivo.*

Assim, denotaremos por $P(A|B)$ *a probabilidade de ter câncer dado que o teste deu positivo, que é justamente o que queremos calcular.*

Do primeiro item temos que $P(A) = 1\%$*, logo,* $P(A^C) = 99\%$*.*

Do segundo item, temos que a probabilidade do teste ser positivo dado que se tem câncer é 80%*, ou seja,* $P(B|A) = 80\%$*.*

A probabilidade de testar positivo dado que se não tem câncer, pelo terceiro item, é $9,6\%$*, isto é,* $P(B|A^C) = 9,6\%$*.*

Usando a ***forma alternativa do teorema de Bayes*** *mencionada anteriormente:*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^C)P(A^C)}$$

$$\Rightarrow P(A|B) = \frac{0,80 \cdot 0,01}{0,80 \cdot 0,01 + 0,096 \cdot 0,99} \approx 0,07764 = 7,764\%.$$

Ou seja, a probabilidade de se ter câncer dado que o teste deu positivo é de *aproximadamente* $7,764\%$.

O exemplo a cima já nos mostra algumas coisas interessantes sobre o Teorema de Bayes.

Embora se pense que o resultado positivo pudesse ser quase que uma certeza de se ter a doença, quando fizemos o cálculo da probabilidade temos que a probabilidade da pessoa ter a doença é bem pequena. Por outro lado, antes do teste a probabilidade de estar doente era de apenas 1%, como deu positivo, essa probabilidade quase que foi multiplicada por 8.

Capítulo 3

Mais aplicações

Nas ultimas décadas, surgiu novo ramo de pesquisa acadêmica que tenta explicar e entender o modo como as pessoas fazem julgamentos e tomam decisões quando defrontadas com informações imperfeitas ou incompletas, tema este muito presente neste trabalho. Em situações que envolvem o acaso, nossos processos cerebrais costumam ser deficientes. Essa é uma área que reúne muitas disciplinas, não só a matemática e as ciências tradicionais, como também a psicologia cognitiva, a economia comportamental e a neurociência moderna. Embora um de seus principais autores, Daniel Kahneman (autor de Rápido e devagar: Duas formas de pensar [4]), tenha ganhado Prêmio Nobel de Economia, em 2002, suas lições, em grande parte, ainda não são conhecidos do grande público. Este capítulo mostra como o Teorema de Bayes pode servir de ferramenta na tomada de decisões, diminuindo a nossa deficiência em relação ao aleatório. Em livros como O Andar do Bebado: Como o Acaso Determina Nossa Vida ([1]), que está na lista dos mais vendidos de Não ficção no Brasil e em Iludidos pelo Acaso: A influência da sorte nos mercados e na vida [2], do ”guru financeiro” Nassim Nicholas Taleb, vemos uma grande tentativa de popularizar a Teoria da Aleatoriedade. Trata dos princípios que governam o acaso, do desenvolvimento dessas ideias e da maneira pela qual elas atuam em política, negócios, medicina, economia, esportes, lazer e outras áreas da atividade humana. Também trata do modo como tomamos decisões e dos processos que nos levam a julgamentos equivocados e decisões ruins quando confrontados com a aleatoriedade ou a incerteza. Ou como essa aleatoriedade muitas vezes nos ajuda, mas não percebemos e achamos que nossas conquistas foram puro mérito nosso, o que poderia até nos levar a um debate sobre a meritocracia, que não é o objetivo deste trabalho. A seguir, trataremos de aplicações da Teoria da Aleatoriedade, mais especificamente, do Teorema de Bayes, usando uma linguagem mais informal, que se alinha com as ideias abordadas nesses livros. A proposta aqui é que a forma leve de tratar aplicações relevantes seja um incentivo para que os alunos estudem o assunto.

3.1 Julgamentos bayesianos

”Todos nós fazemos julgamentos bayesianos”([1]). O filme ”Dança Comigo?”, de 2004, apresenta a confusão que a mulher do protagonista faz entre a probabilidade de que o marido mentisse se estivesse tendo um caso e a probabilidade de que ele estivesse tendo um caso se mentisse. Acontece que o marido começa a tomar aulas de dança escondido, mas sua esposa acredita que sua demora em chegar após o trabalho só pode significar que esteja sendo traída. Então, ela calcula que a chance de que ele esteja mentindo em relação às atividades que realiza após o trabalho é muito maior se ele estiver tendo um caso com outra mulher do que se não estiver, e assim conclui que ele está tendo um caso. Uma situação que mostra claramente o quanto é necessário se tomar cuidado com a ordem dos pensamentos.

Na verdade, esse é um erro bastante comum e é a fonte de força de muitas teorias conspiratórias. Por exemplo, é fácil confundir a ”probabilidade de que haja corrupção no governo dado que ele se enquadra em ideologia de esquerda”com a ”probabilidade de que um governo que se enquadre em ideologia de esquerda se houver corrupção nele”. Sendo essa confusão a origem de consequências mais danosas do que a situação descrita no filme.

Outra fonte de erros é a omissão da conjunção condicional. Vejamos o exemplo, a seguir: Numa família com duas crianças, qual é a probabilidade de que, se uma delas for menina, ambas sejam meninas? Em um primeiro momento, pode-se pensar que a probabilidade de que ambos os bebês fossem meninas seria de 1/4. Isso estaria correto, caso não houvesse a condicional ”se”. Neste caso, estamos preocupados em saber a probabilidade de que as duas crianças sejam meninas dado que uma delas é menina. Já tendo essa informação, o espaço amostral passa a ser o seguinte: (menino,menina),

(menina, menino), (menina, menina). A possibilidade (menino, menino) não faz parte do espaço amostral em questão. Então, a probabilidade passa a ser de 1/3. Ou seja, 1 das 3 opções corresponde ao resultado de que ambas são meninas.

Faz alguns anos, houve uma reportagem que falava sobre nomes curiosos. Um exemplo era o dos filhos que receberam os nomes Xerox, Fotocópia e Autenticada. Suponha que alguma outra família resolvesse imitar a reportagem. Se a família tem duas crianças, qual é a probabilidade, se uma delas for uma menina chamada Fotocópia, de que ambas sejam meninas? Agora informação trata do sexo das crianças e do nome. Nosso espaço amostral original é uma lista de gêneros e nomes. Denotemos "menina chamada Fotocópia" por menina-F e "menina não chamada Fotocópia" por menina-NF, e representemos o espaço amostral desta maneira: (menino, menino), (menino, menina-F), (menino, menina-NF), (menina-F, menino), (menina-NF, menino), (menina-NF, menina-F), (menina-F, menina-NF), (menina-NF, menina-NF) e (menina-F, menina-F). Agora, vamos podar o espaço amostral. Como sabemos que uma das crianças é uma menina chamada Fotocópia, podemos reduzir o espaço amostral a (menino, menina-F), (menina-F, menino), (menina-NF, menina-F), (menina-F, menina-NF) e (menina-F, menina-F). Observe outra diferença entre este problema e o das duas filhas. Aqui, como não é igualmente provável que uma menina se chame ou não Fotocópia, nem todos os elementos do espaço amostral têm a mesma probabilidade.

Portanto, podemos considerar que a chance de que as duas meninas se chamem Fotocópia (embora não possamos afirmar de que se trata de um evento impossível), é tão pequena que podemos desprezar essa possibilidade. Isso nos deixa com apenas (menino, menina-F), (menina-F, menino), (menina-NF, menina-F) e (menina-F, menina-NF), que têm probabilidades extremamente próximas. Como 2 dos 4 elementos do espaço amostral são formados por famílias com duas meninas, a resposta não é mais 1/3 – como era no problema das duas filhas –, e sim 1/2. A informação adicional – o conhecimento sobre o nome da menina – faz diferença.

A teoria de Bayes nos mostra que a probabilidade de que A ocorra se B ocorrer geralmente difere da probabilidade de que B ocorra se A ocorrer. Não levar esse fato em consideração é um erro comum na profissão médica. Por exemplo, em estudos feitos na Alemanha e nos Estados Unidos [7], pesquisadores pediram a médicos que estimassem a probabilidade de que uma mulher assintomática com idade entre 40 e 50 anos que apresentasse uma mamografia positiva realmente tivesse câncer, sabendo que 7% das mamografias mostram câncer quando ele não existe. Além disso, os médicos receberam a informação de que a incidência real da doença era de aproximadamente 0,8%, e que a taxa de falsos negativos era de aproximadamente 10%. Juntando todas as informações, podemos usar o método de Bayes para determinar que uma mamografia positiva representa a presença de câncer em apenas cerca de 9% dos casos. No grupo alemão, no entanto, um terço dos médicos concluiu que a probabilidade seria de 90%, e a mediana das estimativas foi de 70%. No grupo americano, 95 de cada 100 médicos estimaram que a probabilidade seria próxima de 75%.

Outro exemplo de consequências desastrosas a partir do erro da inversão pode ser observado no caso de Sally Clark, na Grã-Bretanha [8]. O primeiro filho de Sally morreu com 11 semanas de idade, em Dezembro de 1996. A causa da morte foi declarada como síndrome da morte súbita infantil (SMSI). Em outras palavras, o bebê morreu inesperadamente e a autópsia não revelou nenhuma causa de morte. Menos de dois anos depois, em janeiro de 1998, Sally ficou grávida novamente. Novamente, ela teve que encarar a morte de um bebê. Dessa vez, às 8 semanas de vida, novamente por SMSI. Dessa vez, a mãe foi presa e acusada de sufocar seus dois filhos. Durante o julgamento, a acusação requisitou o testemunho de um pediatra especialista, sir Roy Meadow. Com base na raridade do diagnóstico, sir Meadow afirmou que a probabilidade de que a causa da morte das duas crianças fosse SMSI era de 1/73 milhões. A acusação não apresentou nenhuma outra prova substancial contra ela. Mas o júri achou que tinha o suficiente para condená-la e, em novembro de 1999, Sally foi mandada para a prisão. O raciocínio do pediatra foi usar uma estimativa para a probabilidade de que uma criança morresse de SMSI, 1/8543. Como eram duas crianças, ele elevou essa estimativa ao quadrado, obtendo 1/73 milhões. No entanto, esse cálculo presume que as mortes sejam independentes. Supõe-se que não haja nenhum efeito ambiental ou genético envolvido, de modo a aumentar o risco do segundo bebê uma vez que seu irmão mais velho tenha morrido de SMSI.

Aqui ocorre novamente o erro da inversão: o que buscamos não é a probabilidade de que duas crianças morram de SMSI, e sim a probabilidade de que as duas crianças que morreram tenham morrido de SMSI. Usando as estatísticas de crimes do Reino Unido e o mesmo raciocínio de Sir Meadows, encontra-se que a probabilidade de duas crianças serem assassinadas na mesma casa é apenas 1/2 bilhões. Baseado nisso, a Royal Statistical Society lançou um comunicado de imprensa no qual declarava que a decisão do júri se baseara em “um grave erro de lógica”. O júri precisa considerar duas explicações concorrentes para as mortes dos bebês: SMSI ou assassinato. Duas mortes por SMSI ou duas mortes por assassinato são ambas bastante improváveis, mas, neste caso, uma delas aparentemente aconteceu. O que importa é a probabilidade relativa das mortes. Nesse caso, vê-se que a probabilidade das duas crianças serem assassinadas é muito menor do que a probabilidade de terem morrido por SMSI.

Os Clark recorreram da sentença e, na apelação, contrataram seus próprios estatísticos como testemunhas. A sentença foi mantida, mas eles continuaram a buscar explicações médicas para as mortes e, no processo, descobriram que o patologista que trabalhava para a acusação havia omitido a informação de que o segundo bebê tinha uma infecção bacteriana no momento da morte, uma infecção que poderia tê-la provocado. Com base nessa descoberta, um juiz revogou a sentença, e após quase três anos e meio, Sally Clark foi libertada da prisão. Sir Meadows usou o mesmo argumento estatístico em casos semelhantes. Foi considerado culpado por má conduta profissional em 2005, tendo esse veredito anulado depois.

3.2 Bayes x Covid

Nesta seção revisitaremos o Exemplo 2.0.2 no contexto da Covid-19. A expectativa é tanto usar a Matemática para explicar questões cotidianas quanto usar tais questões para motivar o aluno a aprender Matemática. Já foi dito no Exemplo 2.0.2 que os exames médicos não são totalmente confiáveis. Sendo assim, uma questão que surge naturalmente é: quantas pessoas que testaram negativo para o SARS CoV 2 estão doentes?

No intuito de responder essa questão, estabeleceremos alguma nomenclatura. Primeiro, deseja-se que um exame tenha alta **sensibilidade**, que é a probabilidade de detectar um caso positivo, e alta **especificidade**, que é a probabilidade de detectar um caso negativo. Diz-se que o exame deu um falso negativo quando o resultado é negativo para a doença, mas o paciente está doente. De modo análogo, o exame dá um falso positivo quando o resultado é positivo para a doença, mas o paciente está saudável.

No caso específico de detecção do SARS CoV 2, a ANVISA tem resgistrados 34 testes do tipo RT-PCR (O nome PCR-RT vem do inglês e significa "reação de transcriptase reversa seguida de reação em cadeia da polimerase"). Desses, 33 apresentam sensibilidade acima de 95% ([5]). Em um teste com 95% de sensibilidade, é de se esperar que em 5% dos doentes o teste dê negativo. Uma confusão frequente é achar que 5 a cada 100 indivíduos que fazem o teste tenham resultados negativos. Note que a sensibilidade é obtida aplicando-se o teste em pacientes infectados. Então temos a probabilidade do exame apresentar resultado positivo em um paciente que se sabe estar doente. Para se fazer uma análise sobre o falso negativo, o que se quer é a probabilidade de o indivíduo estar doente caso o exame resulte em negativo. Para se obter esse valor, precisamos da prevalência da doença na população, que é a sua proporção de doentes. Essa é uma medida que requer um estudo mais complexo sobre a população. Mas façamos a estimativa baseada nos dados do Ministério da Saúde. Segundo os dados disponibilizados em sua página oficial [6], a proporção de casos é de aproximadamente 10 mil para cada 100 mil habitantes, o que significa uma prevalência aproximada de 10%. Há de se chamar a atenção para o fato deste ser um exemplo baseado em aproximações. Também há o fato de se utilizar uma proporção baseada na observação de casos notificados, número que muda todos os dias e está sujeito a erros, como subnotificação, por exemplo.

Assumindo *sensibilidade* de 95%, *especificidade* de 97% e *prevalência* de 10%, dados compatíveis com os exames rT-PCR apresentados em [5], vamos utilizar o Teorema de Bayes para calcular a probabilidade de uma pessoa estar doente mesmo testando negativo.

Definamos como A o evento "tem covid"e B o evento "testou positivo".

Como a *sensibilidade* é a probabilidade de resultado positivo dos doentes, isto é, $P(B|A)$, temos que $P(B|A) = 0,95$. Sendo a especificidade a probabilidade do resultado negativo dos não-doentes, temos que $P(B^C|A^C) = 0,97$. Já a prevalência é a proporção de casos existentes, no nosso caso $P(A)$, logo, $P(A) = 0,10$.

Usando a forma alternativa do Teorema de Bayes 2, semelhante à usada no Exemplo 2.0.2,

$$P(A|B^C) = \frac{P(B^C|A) \cdot P(A)}{P(B^C|A) \cdot P(A) + P(B^C|A^C) \cdot P(A^C)}$$

$$\Rightarrow P(A|B^C) = \frac{0,05 \cdot 0,10}{0,05 \cdot 0,10 + 0,97 \cdot 0,90} \approx 0,0057 = 0,57\%.$$

Assim, mesmo testando negativo, ainda temos $0,57\%$ de probabilidade de estarmos doentes! Pode parecer pouco a primeira vista, mas quando pensamos que essa é a probabilidade de botarmos a vida dos que amamos em risco, ainda parece pouco? Afinal, a cada 200 doentes 1 deve ser falso-negativo. Não abandonem os cuidados sanitários, todo cuidado é pouco.

3.3 Combatendo a pandemia com mais eficácia

E as aplicações do Teorema de Bayes podem ter passado na sua frente e você não percebeu.

Algum tempo atrás quando eu tinha de entrar em algum lugar, era bastante comum ter alguém medindo a temperatura de todos que lá entravam, juntamente com a disponibilidade de álcool em gel e esse tipo de coisa. Para minha surpresa, nos últimos dias não vejo mais termômetros em nenhuma entrada, o que será que aconteceu?

É o Teorema de Bayes nos ajudando novamente! Estudos revelaram que a prática de medir a temperatura na porta não é eficiente para prevenir de pessoas doentes entrarem e que a sensibilidade do teste era baixa e a especificidade era alta, fazendo com que pessoas que estavam com febre por outros motivos, diferentes do covid, fossem impedidas de entrar. Estes estudos podem ser encontrados em [9] e [10].

Conclusão

Por fim, fica evidente que o **Teorema de Bayes**, assim como a Probabilidade, é uma ferramenta poderosa e importante para nossas vidas. Porém, precisamos utiliza-las corretamente para que funcionem bem e tenham o resultado desejado. Para isso é necessário cautela, atenção e conhecimento.

É de **extrema** importância saber quais hipóteses são necessárias para utilizarmos estes conhecimentos de maneira segura e eficaz. Desta maneira, espero que este trabalho traga luz a esta possível lacuna existente e que ao menos abra caminhos para estudos mais profundos.

Referências Bibliográficas

[1] MLODINOW, Leonard. **O andar do bêbado: como o acaso determina nossas vidas**. Editora Schwarcz-Companhia das Letras, 2009.

[2] TALEB, Nassim Nicholas. **Iludidos pelo acaso: A influência da sorte nos mercados e na vida**. Editora Objetiva, 2019.

[3] TURKMAN, Antónia et al. **A falácia do procurador e a COVID-19** http://ceaul.org/a-falacia-do-procurador-e-a-covid-19/ acessado em:11/08/2021.

[4] KAHNEMAN, Daniel. **Rápido e devagar: Duas formas de pensar**. Objetiva, 2012.

[5] VEROTTI, M.P. et al. **Testes diagnósticos para COVID-19 registrados na Agência Nacional de Vigilância Sanitária: sensibilidade e especificidade reportadas pelos fabricantes**. Comunicação em Ciências da Saúde 2020;31 Suppl 1:217-229

[6] https://covid.saude.gov.br/ Acesso em 13/08/2021.

[7] GIGERENZER, Gerd. **Calculated Risks: How to Know When Numbers Deceive You.** Simon and Schuster, 2002.

[8] JOHN Batt. **Stolen Innocence.** Londres, Ebury Press, 2005.

[9] PANĂ, Bogdan C. et al. Real-World Evidence: The Low Validity of Temperature Screening for COVID-19 Triage. Frontiers in public health, v. 9, p. 891, 2021.

[10] BWIRE, George M.; PAULO, Linda S. Coronavirus disease-2019: is fever an adequate screening for the returning travelers?. Tropical medicine and health, v. 48, n. 1, p. 1-3, 2020.

[11] XAVIER, Claudio; BARRETO, Benigno. Matemática aula por aula. São Paulo: FTD, 2005.