

Probabilidade Condicional e Teorema de Bayes

Aplicando Probabilidade Condicional e Teorema de Bayes em Problemas Computacionais

Márcio Nicolau

2025-08-27

Table of contents

Introdução e Objetivos	2
Objetivos de Aprendizagem	2
Probabilidade Condicional	3
Definição e Fórmula	3
Diagrama de Venn	3
Exemplo: Lançamento de Dado	3
Código Probabilidade Condicional	3
Python	3
R	5
Regra da Multiplicação (Probabilidade da Interseção)	6
Exemplo: Seleção de Cartas sem Reposição	6
Código Regra da Multiplicação	6
Python	6
R	7
Eventos Independentes	7
Definição	7
Exemplo: Lançamento de Duas Moedas	8
Lei da Probabilidade Total	8
Diagrama de Partição do Espaço Amostral	8
Exemplo: Linhas de Produção	8
Código Lei da Probabilidade Total	10
Python	10
R	10
Teorema de Bayes	10
Definição e Fórmula	11
Terminologia de Bayes	11

Diagrama de Relações	11
Exemplo: Teste de Doença	11
Código Teorema de Bayes	13
Python	13
R	13
Relação entre os Conceitos	14
Verificação de Aprendizagem	14
Referências Bibliográficas	16

List of Figures

1	Diagrama de Venn para a probabilidade condicional	4
2	Diagrama de Partição do Espaço Amostral	9
3	Diagrama de Relações do Teorema de Bayes	12
4	Diagrama de Fluxo para a relação entre os Conceitos	15

Introdução e Objetivos

Nas aulas anteriores, exploramos os fundamentos da Análise Combinatória e os conceitos básicos de probabilidade, incluindo espaço amostral e eventos. Agora, daremos um passo adiante para entender como a ocorrência de um evento pode influenciar a probabilidade de outro. A **probabilidade condicional** e o **Teorema de Bayes** são ferramentas poderosas que nos permitem atualizar nossas crenças sobre a chance de um evento com base em novas informações ou evidências.

Estes conceitos são pilares na inferência estatística, em sistemas de recomendação, diagnóstico médico, filtragem de spam e muitas outras aplicações na ciência de dados. Eles permitem construir modelos mais realistas e tomar decisões mais informadas em face da incerteza.

Objetivos de Aprendizagem

Ao final desta aula, você será capaz de:

- Definir e calcular a probabilidade condicional de um evento.
- Compreender e aplicar a Regra da Multiplicação para a probabilidade de interseção de eventos.
- Identificar e testar a independência de eventos.
- Aplicar a Lei da Probabilidade Total para calcular probabilidades marginais.
- Entender a intuição e a fórmula do Teorema de Bayes.
- Aplicar o Teorema de Bayes para resolver problemas computacionais, atualizando probabilidades com base em novas evidências.
- Utilizar Python e R para implementar cálculos envolvendo probabilidade condicional e Teorema de Bayes.

Probabilidade Condicional

A **probabilidade condicional** de um evento A dado que outro evento B já ocorreu (ou é sabido que ocorreu) é a probabilidade de A ocorrer sob a nova restrição imposta por B . É denotada por $P(A|B)$, lida como “probabilidade de A dado B ”.

Definição e Fórmula

A probabilidade condicional é definida por: (Bussab; Morettin, 2017, p. 65)

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

onde $P(B) > 0$.

- $P(A \cap B)$ é a probabilidade da interseção de A e B (ou seja, a probabilidade de A e B ocorrerem juntos).
- $P(B)$ é a probabilidade marginal do evento B .

Intuição: Quando sabemos que o evento B ocorreu, nosso espaço amostral efetivo se reduz ao evento B . Assim, $P(A|B)$ mede a proporção de B que também é A .

Diagrama de Venn

Na probabilidade condicional $P(A|B)$, o novo “universo” é o círculo B . Estamos interessados na parte de A que está dentro de B (a interseção $A \cap B$), em proporção ao tamanho total de B .

Exemplo: Lançamento de Dado

Considere o lançamento de um dado justo. $\Omega = \{1, 2, 3, 4, 5, 6\}$.

- Evento A : “Obter um número par” = $\{2, 4, 6\}$. $P(A) = 3/6 = 1/2$.
- Evento B : “Obter um número maior que 3” = $\{4, 5, 6\}$. $P(B) = 3/6 = 1/2$.

Queremos calcular $P(A|B)$: a probabilidade de obter um número par, dado que o número foi maior que 3.

1. **Encontre $A \cap B$:** Resultados que são pares E maiores que 3 = $\{4, 6\}$. $P(A \cap B) = 2/6 = 1/3$.
2. **Calcule $P(A|B)$:** $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/3}{1/2} = \frac{1}{3} \times \frac{2}{1} = \frac{2}{3}$.

Código Probabilidade Condicional

Python

```
# Dados do problema (probabilidades já calculadas ou conhecidas)
p_a = 3/6 # P(Par)
p_b = 3/6 # P(Maior que 3)
p_a_inter_b = 2/6 # P(Par e Maior que 3)

# Verifica se P(B) é maior que 0 para evitar divisão por zero
if p_b > 0:
    p_a_dado_b = p_a_inter_b / p_b
```

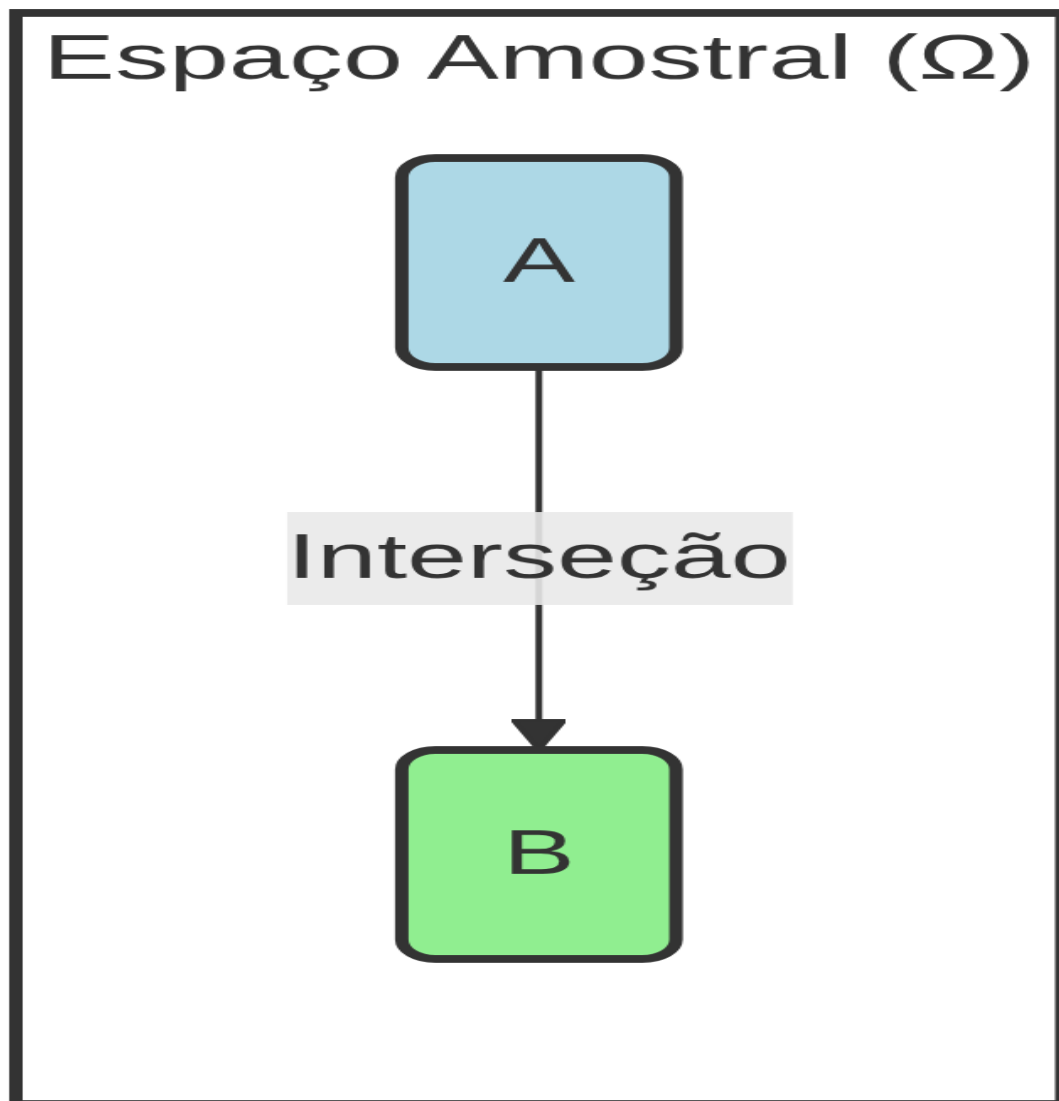


Figure 1: Diagrama de Venn para a probabilidade condicional

```

    print(f"P(A|B) (Probabilidade de ser par dado que é maior que 3): {p_a_dado_b:.3f}")
else:
    print("P(B) é zero, probabilidade condicional indefinida.")

# Exemplo com enumeração de espaço amostral (mais prático para problemas computacionais)
espaco_amostral =
event_A = [x for x in espaco_amostral if x % 2 == 0] # Números pares
event_B = [x for x in espaco_amostral if x > 3]      # Números maiores que 3

# Encontrar a interseção
event_A_inter_B = [x for x in event_A if x in event_B]

# Calcular as probabilidades a partir da contagem
N = len(espaco_amostral)
P_A_inter_B_count = len(event_A_inter_B) / N
P_B_count = len(event_B) / N

if P_B_count > 0:
    P_A_dado_B_count = P_A_inter_B_count / P_B_count
    print(f"P(A|B) (calculado por contagem): {P_A_dado_B_count:.3f}")
else:
    print("P(B) é zero, probabilidade condicional indefinida.")

```

R

```

# Dados do problema (probabilidades já calculadas ou conhecidas)
p_a <- 3/6 # P(Par)
p_b <- 3/6 # P(Maior que 3)
p_a_inter_b <- 2/6 # P(Par e Maior que 3)

# Verifica se P(B) é maior que 0
if (p_b > 0) {
    p_a_dado_b <- p_a_inter_b / p_b
    cat(sprintf("P(A|B) (Probabilidade de ser par dado que é maior que 3): %.3f\n", p_a_dado_b))
} else {
    cat("P(B) é zero, probabilidade condicional indefinida.\n")
}

# Exemplo com enumeração de espaço amostral
espaco_amostral <- 1:6
event_A <- espaco_amostral[espaco_amostral %% 2 == 0] # Números pares
event_B <- espaco_amostral[espaco_amostral > 3]      # Números maiores que 3

# Encontrar a interseção

```

```

event_A_inter_B <- intersect(event_A, event_B)

# Calcular as probabilidades a partir da contagem
N <- length(espaco_amostrai)
P_A_inter_B_count <- length(event_A_inter_B) / N
P_B_count <- length(event_B) / N

if (P_B_count > 0) {
  P_A_dado_B_count <- P_A_inter_B_count / P_B_count
  cat(sprintf("P(A|B) (calculado por contagem): %.3f\n", P_A_dado_B_count))
} else {
  cat("P(B) é zero, probabilidade condicional indefinida.\n")
}

```

Regra da Multiplicação (Probabilidade da Interseção)

A fórmula da probabilidade condicional pode ser rearranjada para obter a **Regra da Multiplicação**, que nos permite calcular a probabilidade da interseção de dois eventos. (Bussab; Morettin, 2017, p. 67)

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

$$\text{Alternativamente, } P(A \cap B) = P(B|A)P(A)$$

Esta regra pode ser estendida para mais de dois eventos. Por exemplo, para três eventos A, B, C : $P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)$.

Exemplo: Seleção de Cartas sem Reposição

Um baralho padrão de 52 cartas. Duas cartas são retiradas *sem reposição*. Qual a probabilidade de que ambas as cartas sejam Ases?

- Evento A: “A primeira carta é um Ás”. $P(A) = 4/52$.
- Evento B: “A segunda carta é um Ás”, *dado que* a primeira foi um Ás. Após retirar um Ás, restam 51 cartas no baralho e 3 Ases. $P(B|A) = 3/51$.

A probabilidade de ambas serem Ases é $P(A \cap B)$:

$$P(A \cap B) = P(A)P(B|A) = \frac{4}{52} \times \frac{3}{51} = \frac{12}{2652} = \frac{1}{221} \approx 0.0045.$$

Código Regra da Multiplicação

Python

```

# Dados do problema
num_ases = 4
total_cartas = 52

# P(A) = Probabilidade da primeira carta ser um Ás

```

```

p_primeiro_as = num_ases / total_cartas
print(f"P(Primeira carta é Ás): {p_primeiro_as:.3f}")

# P(B|A) = Probabilidade da segunda carta ser um Ás, dado que a primeira foi um Ás
p_segundo_as_dado_primeiro_as = (num_ases - 1) / (total_cartas - 1)
print(f"P(Segunda carta é Ás | Primeira carta foi Ás): {p_segundo_as_dado_primeiro_as:.3f}")

# P(A e B) = P(Ambas as cartas são Ases)
p_ambos_ases = p_primeiro_as * p_segundo_as_dado_primeiro_as
print(f"P(Ambas as cartas são Ases): {p_ambos_ases:.4f}")

```

R

```

# Dados do problema
num_ases <- 4
total_cartas <- 52

# P(A) = Probabilidade da primeira carta ser um Ás
p_primeiro_as <- num_ases / total_cartas
cat(sprintf("P(Primeira carta é Ás): %.3f\n", p_primeiro_as))

# P(B|A) = Probabilidade da segunda carta ser um Ás, dado que a primeira foi um Ás
p_segundo_as_dado_primeiro_as <- (num_ases - 1) / (total_cartas - 1)
cat(sprintf("P(Segunda carta é Ás | Primeira carta foi Ás): %.3f\n", p_segundo_as_dado_primeiro_as))

# P(A e B) = P(Ambas as cartas são Ases)
p_ambos_ases <- p_primeiro_as * p_segundo_as_dado_primeiro_as
cat(sprintf("P(Ambas as cartas são Ases): %.4f\n", p_ambos_ases))

```

Eventos Independentes

Dois eventos A e B são considerados **independentes** se a ocorrência de um não afeta a probabilidade de ocorrência do outro. (Bussab; Morettin, 2017, p. 69)

Definição

Matematicamente, A e B são independentes se:

- $P(A|B) = P(A)$ (a probabilidade de A não muda, mesmo sabendo que B ocorreu).
- Ou, $P(B|A) = P(B)$ (a probabilidade de B não muda, mesmo sabendo que A ocorreu).

A partir da Regra da Multiplicação, se A e B são independentes, a probabilidade da interseção é simplesmente o produto das probabilidades marginais:

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

Se essa condição não for satisfeita, os eventos são **dependentes**.

Exemplo: Lançamento de Duas Moedas

Considere o lançamento de duas moedas justas. $\Omega = \{CC, CK, KC, KK\}$.

- Evento A: “A primeira moeda resulta em Cara” = $\{CC, CK\}$. $P(A) = 2/4 = 1/2$.
- Evento B: “A segunda moeda resulta em Cara” = $\{CC, KC\}$. $P(B) = 2/4 = 1/2$.
- $A \cap B$: “Ambas as moedas resultam em Cara” = $\{CC\}$. $P(A \cap B) = 1/4$.

Para verificar a independência, usamos $P(A \cap B) = P(A)P(B)$:

$1/4 = (1/2) \times (1/2)$ $1/4 = 1/4$. Como a igualdade é verdadeira, os eventos A e B são independentes.

Lei da Probabilidade Total

A **Lei da Probabilidade Total** é uma regra fundamental que permite calcular a probabilidade de um evento A somando as probabilidades de A em cada cenário possível, onde os cenários formam uma partição do espaço amostral. (Bussab; Morettin, 2017, p. 73)

Considere que $B_1, B_2, \dots, B_n$ são eventos mutuamente exclusivos (disjuntos) e coletivamente exaustivos (sua união cobre todo o espaço amostral, $\bigcup_{i=1}^n B_i = \Omega$). Então, a probabilidade de qualquer evento A pode ser calculada como:

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i)$$

Ou, para o caso de dois eventos B e B^c (complemento de B):

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

Diagrama de Partição do Espaço Amostral

(Imagine que o evento A se sobrepõe a partes de B_1 , B_2 e B_3 . A Lei da Probabilidade Total soma a contribuição de A em cada partição B_i para obter $P(A)$.)

Exemplo: Linhas de Produção

Uma empresa possui duas linhas de produção, L1 e L2. L1 produz 60% dos itens e L2 produz 40%. A linha L1 tem uma taxa de defeito de 3%, e a linha L2 tem uma taxa de defeito de 5%. Um item é selecionado aleatoriamente da produção total. Qual a probabilidade de que o item seja defeituoso?

- Evento D: “O item é defeituoso”.
- Evento L1: “O item veio da linha L1”. $P(L1) = 0.60$.
- Evento L2: “O item veio da linha L2”. $P(L2) = 0.40$. (L1 e L2 formam uma partição do espaço amostral).
- Probabilidade de defeito, dado que veio de L1: $P(D|L1) = 0.03$.
- Probabilidade de defeito, dado que veio de L2: $P(D|L2) = 0.05$.

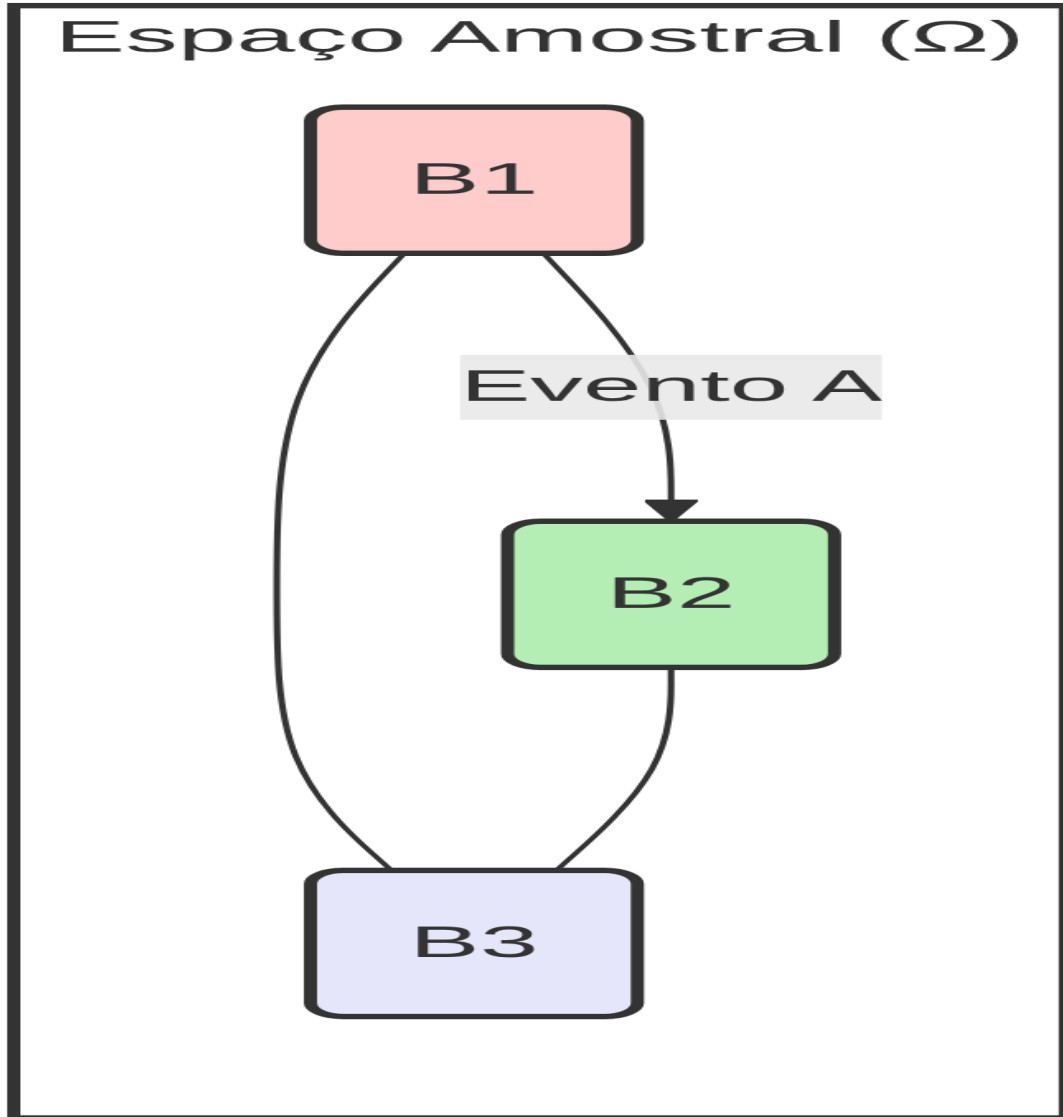


Figure 2: Diagrama de Partição do Espaço Amostral

Usando a Lei da Probabilidade Total:

- $P(D) = P(D|L1)P(L1) + P(D|L2)P(L2)$
- $P(D) = (0.03)(0.60) + (0.05)(0.40)$
- $P(D) = 0.018 + 0.020$
- $P(D) = 0.038$

A probabilidade de um item selecionado aleatoriamente ser defeituoso é de 3.8%.

Código Lei da Probabilidade Total

Python

```
# Probabilidades de cada linha de produção
p_l1 = 0.60
p_l2 = 0.40

# Probabilidades de defeito dado cada linha
p_d_dado_l1 = 0.03
p_d_dado_l2 = 0.05

# Lei da Probabilidade Total
p_d = (p_d_dado_l1 * p_l1) + (p_d_dado_l2 * p_l2)

print(f"P(D) (Probabilidade total de um item ser defeituoso): {p_d:.3f}")
```

R

```
# Probabilidades de cada linha de produção
p_l1 <- 0.60
p_l2 <- 0.40

# Probabilidades de defeito dado cada linha
p_d_dado_l1 <- 0.03
p_d_dado_l2 <- 0.05

# Lei da Probabilidade Total
p_d <- (p_d_dado_l1 * p_l1) + (p_d_dado_l2 * p_l2)

cat(sprintf("P(D) (Probabilidade total de um item ser defeituoso): %.3f\n", p_d))
```

Teorema de Bayes

O **Teorema de Bayes** é uma fórmula que nos permite atualizar a probabilidade de uma hipótese (ou causa) dado que observamos uma evidência (ou efeito). Ele relaciona a probabilidade condicional de A dado B com

a probabilidade condicional de B dado A . (Bussab; Morettin, 2017, p. 73–74)

Definição e Fórmula

Dada a probabilidade condicional $P(A|B) = \frac{P(A \cap B)}{P(B)}$ e a Regra da Multiplicação $P(A \cap B) = P(B|A)P(A)$, podemos substituir $P(A \cap B)$ na primeira fórmula:

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

Aqui, $P(B)$ no denominador pode ser expandido usando a Lei da Probabilidade Total (se A_i formam uma partição do espaço amostral):

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

Terminologia de Bayes

- $P(A)$: **Probabilidade a priori (Prior)** - Nossa crença inicial na probabilidade de A antes de observar a evidência B .
- $P(B|A)$: **Verossimilhança (Likelihood)** - A probabilidade de observar a evidência B se a hipótese A for verdadeira.
- $P(B)$: **Evidência (Marginal Evidence)** - A probabilidade total de observar a evidência B , independentemente de A . (Normalmente calculada pela Lei da Probabilidade Total).
- $P(A|B)$: **Probabilidade a posteriori (Posterior)** - A probabilidade atualizada de A após observar a evidência B .

Diagrama de Relações

Exemplo: Teste de Doença

Uma doença rara afeta 1% da população ($P(Doença) = 0.01$). Existe um teste para detectar a doença com as seguintes características:

- A taxa de **verdadeiros positivos** (o teste dá positivo se a pessoa tem a doença) é de 95% ($P(\text{Positivo}|Doença) = 0.95$).
- A taxa de **falsos positivos** (o teste dá positivo se a pessoa *não* tem a doença) é de 10% ($P(\text{Positivo}|\text{Não Doença}) = 0.10$).

Se uma pessoa aleatoriamente selecionada testa positivo, qual a probabilidade de que ela realmente tenha a doença? Ou seja, queremos calcular $P(Doença|\text{Positivo})$.

Vamos definir os eventos:

- D : A pessoa tem a doença. $P(D) = 0.01$.
- $\neg D$: A pessoa não tem a doença. $P(\neg D) = 1 - P(D) = 1 - 0.01 = 0.99$.
- P : O teste deu positivo.

Conhecemos:

- $P(P|D) = 0.95$ (verossimilhança)
- $P(P|\neg D) = 0.10$ (falso positivo)

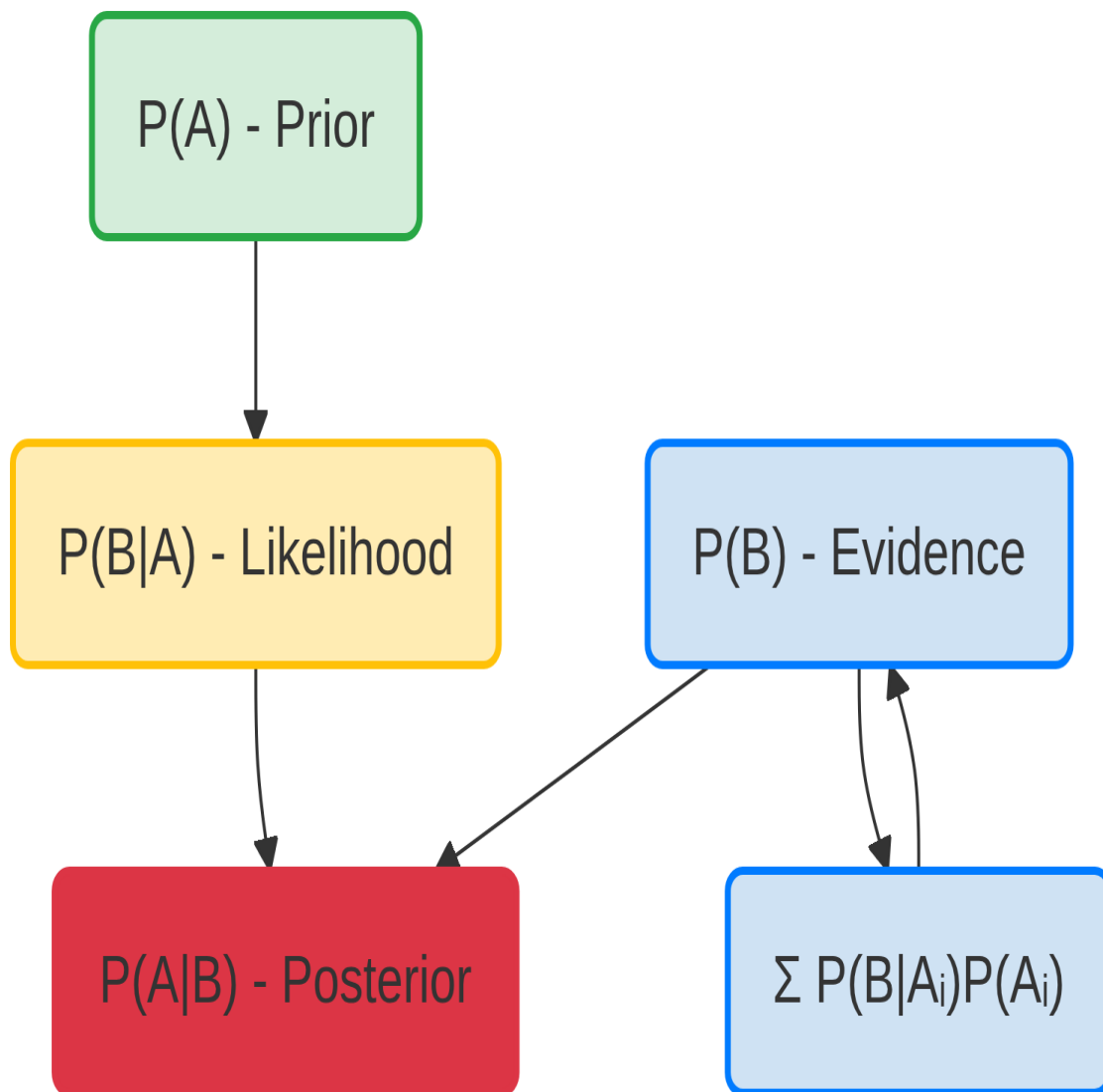


Figure 3: Diagrama de Relações do Teorema de Bayes

Queremos encontrar $P(D|P)$.

Pelo Teorema de Bayes: $P(D|P) = \frac{P(P|D)P(D)}{P(P)}$

Primeiro, precisamos calcular $P(P)$ usando a Lei da Probabilidade Total:

- $P(P) = P(P|D)P(D) + P(P|\neg D)P(\neg D)$
- $P(P) = (0.95)(0.01) + (0.10)(0.99)$
- $P(P) = 0.0095 + 0.0990$
- $P(P) = 0.1085$

Agora, calculamos $P(D|P)$:

$$P(D|P) = \frac{(0.95)(0.01)}{0.1085} = \frac{0.0095}{0.1085} \approx 0.0875$$

Mesmo com um teste positivo, a probabilidade de realmente ter a doença é de apenas cerca de 8.75%! Isso ocorre porque a doença é muito rara na população (prior baixo), e a taxa de falsos positivos, embora pareça pequena, é significativa quando aplicada a uma grande maioria de pessoas saudáveis.

Código Teorema de Bayes

Python

```
# Probabilidades a priori (Prior)
p_doenca = 0.01
p_nao_doenca = 1 - p_doenca

# Verossimilhanças (Likelihood)
p_pos_dado_doenca = 0.95 # Verdadeiro positivo
p_pos_dado_nao_doenca = 0.10 # Falso positivo

# 1. Calcular a Probabilidade Total da Evidência (P(Positivo))
p_positivo = (p_pos_dado_doenca * p_doenca) + \
    (p_pos_dado_nao_doenca * p_nao_doenca)
print(f"P(Positivo) (Probabilidade total de um teste positivo): {p_positivo:.4f}")

# 2. Aplicar o Teorema de Bayes
p_doenca_dado_pos = (p_pos_dado_doenca * p_doenca) / p_positivo

print(f"P(Doença | Positivo) (Probabilidade de ter a doença dado um teste positivo): {p_doenca_dado_pos:.4f}")
```

R

```
# Probabilidades a priori (Prior)
p_doenca <- 0.01
p_nao_doenca <- 1 - p_doenca

# Verossimilhanças (Likelihood)
```

```

p_pos_dado_doenca <- 0.95 # Verdadeiro positivo
p_pos_dado_ao_doenca <- 0.10 # Falso positivo

# 1. Calcular a Probabilidade Total da Evidência (P(Positivo))
p_positivo <- (p_pos_dado_doenca * p_doenca) +
              (p_pos_dado_ao_doenca * p_ao_doenca)
cat(sprintf("P(Positivo) (Probabilidade total de um teste positivo): %.4f\n", p_positivo))

# 2. Aplicar o Teorema de Bayes
p_doenca_dado_pos <- (p_pos_dado_doenca * p_doenca) / p_positivo

cat(sprintf("P(Doença | Positivo) (Probabilidade de ter a doença dado um teste positivo): %.4f\n", p_doenca_dado_pos))

```

Relação entre os Conceitos

Estes conceitos se constroem mutuamente:

- A **probabilidade condicional** é a base para entender como eventos se relacionam.
- A **Regra da Multiplicação** deriva da probabilidade condicional e permite calcular a probabilidade de eventos conjuntos.
- A noção de **eventos independentes** é um caso especial da probabilidade condicional/regra da multiplicação, onde a ocorrência de um evento não influencia o outro.
- A **Lei da Probabilidade Total** é crucial para calcular a probabilidade marginal de uma evidência que pode ser gerada por diferentes causas (hipóteses).
- O **Teorema de Bayes** usa a probabilidade condicional (na forma de verossimilhança) e a probabilidade total da evidência para inverter a condicionalidade, permitindo-nos atualizar nossas probabilidades a priori em probabilidades a posteriori.

Verificação de Aprendizagem

Resolva os problemas abaixo, aplicando os conceitos de probabilidade condicional, Regra da Multiplicação, Lei da Probabilidade Total e Teorema de Bayes. Utilize código em Python ou R para auxiliar nos cálculos quando apropriado.

1. Pesquisa de Mercado:

Uma pesquisa de mercado mostrou que 30% das pessoas de uma cidade preferem a Marca A de refrigerante, 50% preferem a Marca B e 20% preferem outras marcas. Sabe-se também que, entre as pessoas que preferem a Marca A, 60% são do sexo feminino. Entre as que preferem a Marca B, 40% são do sexo feminino. E entre as que preferem outras marcas, 70% são do sexo feminino.

- a) Qual a probabilidade de uma pessoa selecionada aleatoriamente ser do sexo feminino?
- b) Se uma pessoa selecionada aleatoriamente é do sexo feminino, qual a probabilidade de que ela prefira a Marca A?

2. Qualidade de Peças:

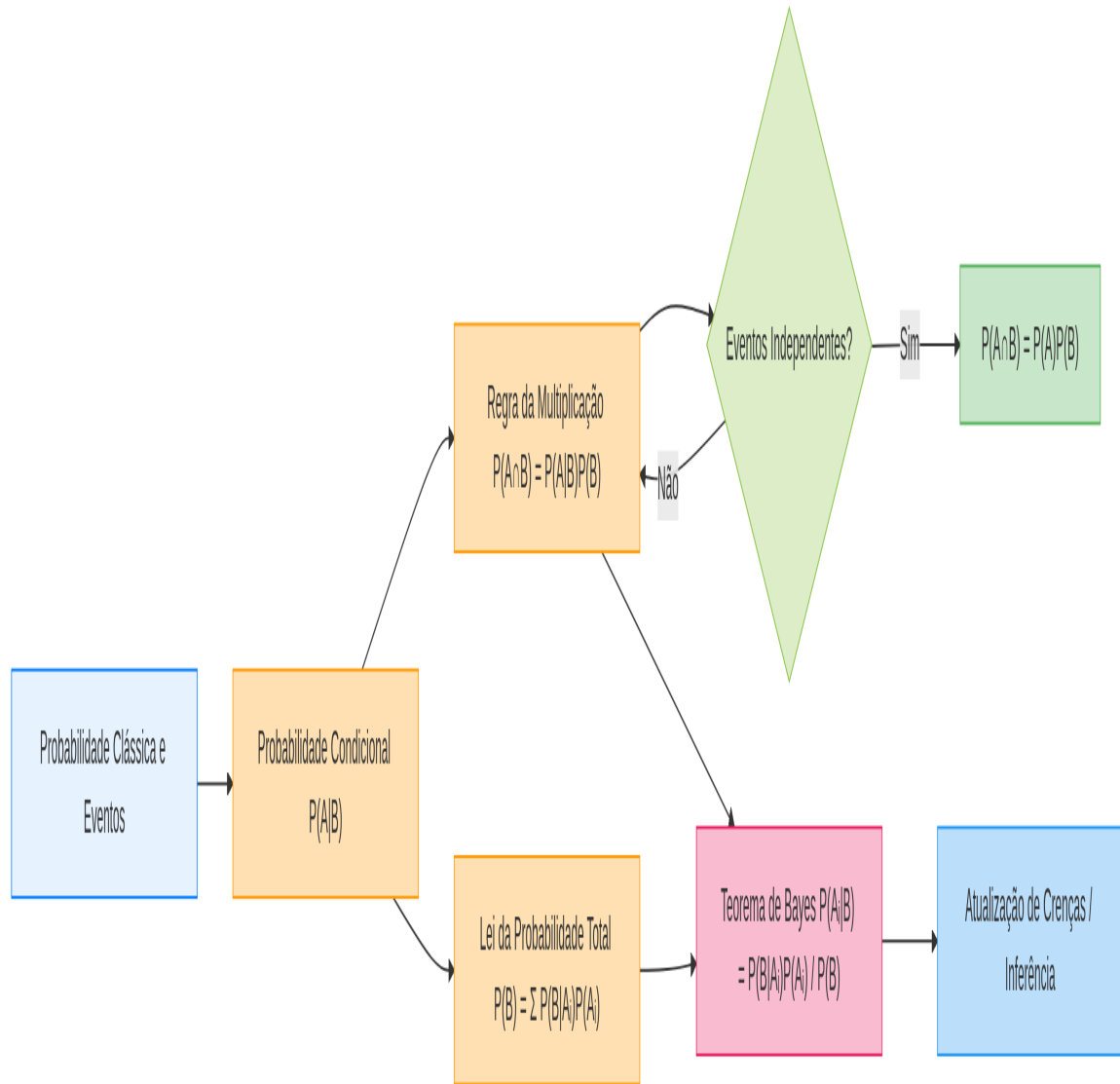


Figure 4: Diagrama de Fluxo para a relação entre os Conceitos

Uma fábrica produz parafusos em três máquinas, M1, M2 e M3.

- M1 produz 40% dos parafusos e tem uma taxa de defeito de 2%.
- M2 produz 35% dos parafusos e tem uma taxa de defeito de 3%.
- M3 produz 25% dos parafusos e tem uma taxa de defeito de 4%.

Um parafuso é selecionado aleatoriamente e é encontrado defeituoso. Qual a probabilidade de que ele tenha sido produzido pela máquina M1?

3. Sintoma de Doença:

A probabilidade de uma pessoa ter uma certa doença é de 0.005. Um teste de diagnóstico para essa doença tem uma sensibilidade (probabilidade de dar positivo dado que a pessoa tem a doença) de 98% e uma especificidade (probabilidade de dar negativo dado que a pessoa *não* tem a doença) de 90%.

- a) Calcule a probabilidade de uma pessoa ter a doença, dado que o teste deu positivo.
- b) Calcule a probabilidade de uma pessoa não ter a doença, dado que o teste deu negativo.

4. Dados e Moeda:

Você tem uma moeda comum e um dado comum. Você joga a moeda. Se der “Cara”, você joga o dado uma vez. Se der “Coroa”, você joga o dado duas vezes e soma os resultados.

- a) Qual a probabilidade de obter um 6 no dado (se jogou uma vez) ou uma soma de 6 (se jogou duas vezes)?
- b) Se você obteve um resultado de 6 (no dado ou na soma), qual a probabilidade de que a moeda tenha dado “Cara”?

Referências Bibliográficas

BUSSAB, Luiz O. de M.; MORETTIN, Pedro A. **Estatística Básica**. 9. ed. São Paulo: Saraiva, 2017.