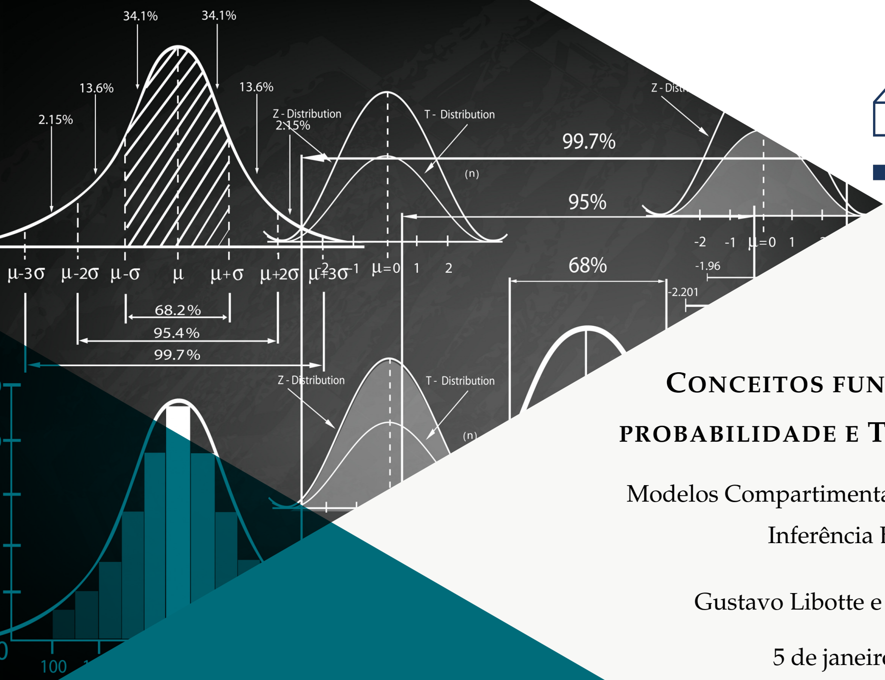


Laboratório Nacional de Computação Científica
UNIDADE DE PESQUISA DO MCTI

Conceitos fundamentais de probabilidade e Teorema de Bayes

Modelos Compartimentais em Epidemiologia e Inferência Bayesiana

Gustavo Libotte e Regina Almeida

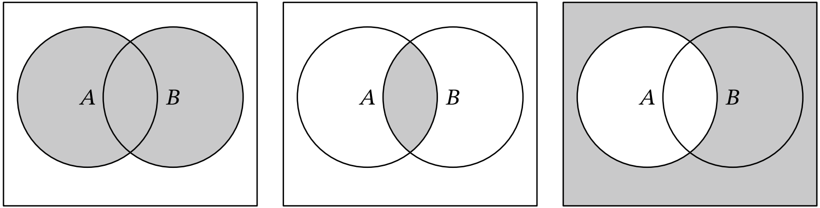
5 de janeiro de 2022

Conceitos fundamentais de probabilidade
Conjuntos

- Um **conjunto** é uma coleção de objetos. Por exemplo, $A = \{1, 2, 3, 11\}$ ou $A = \{x \mid 0 \leq x \leq 1\}$.
- Os objetos que individualmente formam a coleção ou conjunto A são denominados **membros** ou **elementos** de A.
- Quando dois conjuntos A e B são considerados, e ser elemento de A implica em ser elemento de B, então diremos que A é um **subconjunto** de B, e escreveremos $A \subset B$.
- **Para todo** conjunto A, temos que $\emptyset \subset A$.
- Definiremos C como a **união** de A e B (algumas vezes denominada a **soma** de A e B), da seguinte maneira:
$$C = \{x \mid x \in A \text{ ou } x \in B \text{ (ou ambos)}\}$$
- Escreveremos a **união** de A e B como $C = A \cup B$.

Conceitos fundamentais de probabilidade
Conjuntos

- Definiremos D como a **interseção** de A e B (algumas vezes denominado o **produto** de A e B), da seguinte maneira:
$$D = \{x \mid x \in A \text{ e } x \in B\}$$
- Escreveremos a **interseção** de A e B como $D = A \cap B$.
- O conjunto denotado por $\bar{A}$, constituído por todos os elementos que **não** estejam em A é denominado **complemento** de A:
$$\bar{A} = \{x \mid x \notin A\}$$



$A \cup B$ $\qquad$ $A \cap B$ $\qquad$ $\bar{A}$

Conceitos fundamentais de probabilidade
Espaço amostral e eventos

- Para cada experimento ε, definimos o **espaço amostral** como o conjunto de todos os resultados possíveis de ε. Geralmente representaremos esse conjunto por S.
- Por exemplo, para o experimento
$$E_1 : \text{Jogue um dado e observe o número mostrado na face de cima}$$
o espaço amostral correspondente é
$$S_1 : \{1, 2, 3, 4, 5, 6\}$$
- Um **evento** A, relativo a um espaço amostral particular S e associado a um experimento ε, é simplesmente um **conjunto de resultados possíveis**.
- Na terminologia dos conjuntos, um evento é um **subconjunto** de um espaço amostral S.
- O próprio espaço amostral S constitui um evento, bem como o conjunto vazio $\emptyset$. Qualquer resultado individual pode também ser tomado como um evento.
- $A_1 : \{2, 4, 6\}$ é um evento possível para o experimento E_1.

Conceitos fundamentais de probabilidade
Espaço amostral e eventos

- Dois eventos A e B são denominados **mutuamente excludentes** se eles não puderem ocorrer juntos.
- Isso pode ser exprimido escrevendo $A \cap B = \emptyset$, isto é, a interseção de A e B é o **conjunto vazio**.
- Exemplo: Um dispositivo eletrônico é ensaiado e o tempo total de serviço t é registrado. Admitiremos que o espaço amostral seja $\{t \mid t \geq 0\}$. Sejam A, B e C três eventos definidos da seguinte maneira:

  $$A = \{t \mid t < 100\}; \qquad B = \{t \mid 50 \leq t \leq 200\}; \qquad C = \{t \mid t > 150\}$$

  Consequentemente:

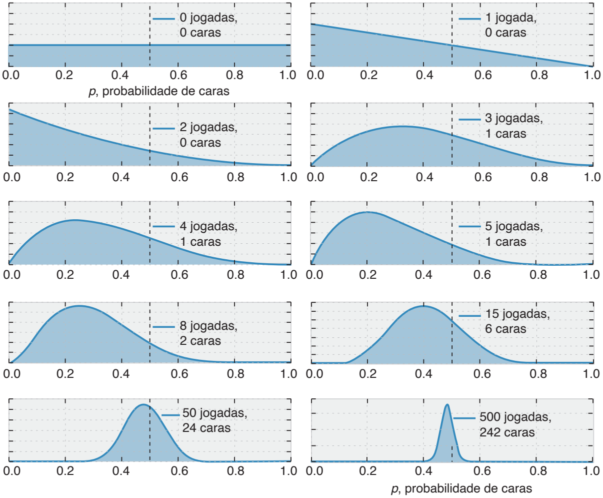
  $$A \cup B = \{t \mid t \leq 200\}; \qquad A \cap B = \{t \mid 50 \leq t < 100\}$$
  $$B \cup C = \{t \mid t \geq 50\}; \qquad B \cap C = \{t \mid 150 < t \leq 200\}; \qquad A \cap C = \emptyset;$$
  $$A \cup C = \{t \mid t < 100 \text{ ou } t > 150\}; \qquad \bar{A} = \{t \mid t \geq 100\}; \qquad \bar{C} = \{t \mid t \leq 150\}.$$

Conceitos fundamentais de probabilidade
Frequência relativa

- Suponha que repetimos n vezes o experimento ε, e sejam A e B dois eventos associados a ε. Admita que sejam, respectivamente, n_A e n_B o número de vezes que os eventos A e B ocorreram nas n repetições.
- $f_A = n_A/n$ é denominada **frequência relativa** do evento A nas n repetições de ε. A frequência relativa f_A apresenta as seguintes propriedades:
  1. $0 \leq f_A \leq 1$.
  2. $f_A = 1$ se, e somente se, A ocorrer em todas as n repetições.
  3. $f_A = 0$ se, e somente se, A nunca ocorrer nas n repetições.
  4. Se A e B forem eventos mutuamente excludentes, e se $f_{A \cup B}$ for a frequência relativa associada ao evento $A \cup B$, então $f_{A \cup B} = f_A + f_B$.
  5. f_A, com base em n repetições do experimento e considerada como uma função de n, "converge" em certo sentido probabilístico para $P(A)$, quando $n \to \infty$.
- Se um experimento for executado um grande número de vezes, a frequência relativa da ocorrência de alguma evento A tenderá a variar cada vez menos à medida que o número de repetições for aumentada. Isso é conhecido como **regularidade estatística**.

Conceitos fundamentais de probabilidade
Frequência relativa

- Exemplo: Começamos a jogar uma **moeda** não-viciada e registramos as observações de cara ou coroa. Este são os nosso dados observados.
- Uma pergunta interessante a fazer é: como nossa inferência de p muda à medida que observamos **mais e mais lançamentos da moeda**?



Atividade prática

Implemente este experimento e mostre a evolução da distribuição de probabilidade.

Conceitos fundamentais de probabilidade
Probabilidade

- Seja ε um experimento e seja S um espaço amostral associado a ε. A cada evento A associaremos um número real representado por $P(A)$ e denominado **probabilidade** de A, que satisfaça as seguintes propriedades:

  1. $0 \leq P(A) \leq 1$.
  2. $P(S) = 1$.
  3. Se A e B forem eventos mutuamente excludentes, $P(A \cup B) = P(A) + P(B)$.

- Observe-se que da Propriedade 3, se $A_1, A_2, \ldots, A_n$ forem, dois a dois, eventos mutuamente excludentes, decorre imediatamente que, para qualquer n finito,

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

- Essas propriedades são conhecidas como os **axiomas de Kolmogorov**.

Conceitos fundamentais de probabilidade
Teoremas importantes

- Se $\emptyset$ for o conjunto vazio, então $P(\emptyset) = 0$.
- Se $\bar{A}$ for o evento complementar de A, então $P(A) = 1 - P\left(\bar{A}\right)$.
- Se A e B forem dois eventos quaisquer, então $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- Se A, B e C forem três eventos quaisquer, então

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)\ .$$

- Se $A \subset B$, então $P(A) \leq P(B)$.

Para ver as demonstrações, consulte [1].

[1] P. L. Meyer.
Probabilidade.
LTC, Rio de Janeiro, 2 edition, 1987.

Conceitos fundamentais de probabilidade
Probabilidade condicionada

Um exemplo motivacional:

- Um lote de 100 peças é composto de 20 peças defeituosas e 80 peças perfeitas. Suponha-se que escolhemos duas peças desse lote: (a) **com reposição**; (b) **sem reposição**.
- Vamos definir os **eventos**:

$$A = \{\text{a primeira peça é defeituosa}\} \qquad B = \{\text{a segunda peça é defeituosa}\}$$

- Se estivermos extraindo **com reposição**:

$$P(A) = P(B) = 20/100 = 1/5$$

  Porque cada vez que extrairmos do lote, **existirão 20 peças defeituosas** no total de 100.
- No entanto, se estivermos extraindo **sem reposição**, os resultados não serão tão imediatos.
- É ainda verdade, naturalmente, que $P(A) = 1/5$. Mas e sobre $P(B)$?
- A fim de calcularmos $P(B)$, deveremos conhecer a composição do lote **no momento de se extrair a segunda peça**. Isto é, deveremos saber se A ocorreu ou não.

Conceitos fundamentais de probabilidade
Probabilidade condicionada

- Sejam A e B dois eventos associados ao experimento ε. Denotaremos por $P(B \mid A)$ a **probabilidade condicionada** do evento B, quando A tiver ocorrido.
- No nosso exemplo, $P(B \mid A) = 19/99$, porque se A tiver ocorrido, então para a segunda extração **restarão somente 99 peças**, das quais **19 delas serão defeituosas**.
- Sempre que calcularmos $P(B \mid A)$, estaremos essencialmente calculando $P(B)$ em relação ao **espaço amostral reduzido** A, ao invés de fazê-lo em relação ao espaço amostral original S.

Veja a diferença

- Quando calcularmos $P(B)$ estaremos nos perguntando quão provável será obtermos B, **considerando** todo o espaço amostral S.
- E quando calcularmos $P(B \mid A)$, estaremos perguntando quão provável será obtermos B, mas agora **considerando** A, isto é, o espaço amostral ficou reduzido de S para A.

Conceitos fundamentais de probabilidade
Probabilidade condicionada

Mais um exemplo:

- Dois dados equilibrados são lançados, registrando-se o resultado como (x_1, x_2), onde x_i é o resultado do i-ésimo dado, para $i = 1, 2$. Por isso, o espaço amostral S pode ser representado pela seguinte lista de **36 resultados igualmente prováveis**.

$$S = \left\{ \begin{array}{cccc} (1,1) & (1,2) & \ldots & (1,6) \\ (2,1) & (2,2) & \ldots & (2,6) \\ \vdots & & & \vdots \\ (6,1) & (6,2) & \ldots & (6,6) \end{array} \right\}$$

- Consideremos os dois **eventos** seguintes:

$$A = \{(x_1, x_2) \mid x_1 + x_2 = 10\} \qquad B = \{(x_1, x_2) \mid x_1 > x_2\}$$

- Assim, $A = \{(5,5), (4,6), (6,4)\}$ e $B = \{(2,1), (3,1), (3,2), \ldots, (6,5)\}$.

Conceitos fundamentais de probabilidade
Probabilidade condicionada

- Portanto, $P(A) = 3/36$, $P(B) = 15/36$ e $P(B \mid A) = 1/3$, uma vez que **o espaço amostral** é, agora, formado por A (isto é, três resultados possíveis), e **somente um** desses três resultados é coerente com o evento B.
- Finalmente, vamos calcular $P\left(A \cap B\right)$.
- O evento $A \cap B$ ocorre **se, e somente se**, a soma dos dois dados for 10 e se o primeiro dado tiver, apresentado um valor maior que o segundo dado.
- Existe apenas **um desses resultados** e, por isso, $P\left(A \cap B\right) = 1/36$.
- Se fizermos um exame cuidadoso dos vários números já calculados, concluiremos que

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{e} \quad P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

Conceitos fundamentais de probabilidade
Probabilidade condicionada

- Admita que um experimento ε tenha sido repetido n vezes. Sejam n_A, n_B e $n_{A\cap B}$ o **número de vezes** que, respectivamente, os eventos A, B e $A \cap B$ tenham ocorrido em n repetições.
- Então, $n_{A\cap B}/n_A$ representa a frequência relativa de B, condicionada a que A tenha ocorrido.
- Poderemos escrever $n_{A\cap B}/n_A$, da seguinte forma:

$$\frac{n_{A\cap B}}{n_A} = \frac{n_{A\cap B}/n}{n_A/n} = \frac{f_{A\cap B}}{f_A}$$

  onde $f_{A\cap B}$ e f_A são as frequências relativas dos eventos $A \cap B$ e A, respectivamente.
- Se n for grande, $f_{A\cap B}$ será próxima de $P(A \cap B)$ e f_A será próxima de $P(A)$.
- Consequentemente, a relação acima sugere que $n_{A\cap B}/n_A$ será próxima de $P(B \mid A)$. Por isso, estabeleceremos a seguinte definição:

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

Conceitos fundamentais de probabilidade
Teorema da multiplicação e partições

- A mais importante consequência da definição de probabilidade condicionada é obtida com
$$P(A \cap B) = P(B \mid A)P(A), \text{ ou equivalentemente, } P(A \cap B) = P(A \mid B)P(B)$$
Isto é, algumas vezes, mencionado como o **teorema da multiplicação** de probabilidades.
- Até aqui, empregamos o conceito de probabilidade condicionada a fim de avaliar **a probabilidade de ocorrência conjunta** de dois eventos.
- Poderemos aplicar esse conceito em outra maneira de calcular a probabilidade de **um evento simples** A. Necessitaremos da seguinte definição:
- Dizemos que os eventos $B_1, B_2, \ldots, B_k$ representam uma **partição** do espaço amostral S, quando
    1. $B_i \cap B_j = \emptyset$, para todo $i \neq j$.
    2. $\bigcup_{i=1}^{k} B_i = S$.
    3. $P(B_i) > 0$ para todo i.
- Explicando: Quando o experimento ε é realizado **um, e somente um**, dos eventos B_i ocorre.

Conceitos fundamentais de probabilidade
Partições e teorema da probabilidade total

- Considere A uma **evento** qualquer referente a S, e $B_1, B_2, \ldots, B_k$ uma **partição** de S. Portanto, podemos escrever
$$A = A \cap B_1 \cup A \cap B_2 \cup \cdots \cup A \cap B_k$$
- Todos os eventos $A \cap B_1, \ldots, A \cap B_k$ são dois a dois mutuamente excludentes (não podem ocorrer simultaneamente).
- Podemos aplicar a propriedade da adição de eventos mutuamente excludentes e escrever
$$P(A) = P\left(A \cap B_1\right) + P\left(A \cap B_2\right) + \cdots + P\left(A \cap B_k\right)$$
- Contudo, cada termo pode ser expresso na forma $P\left(A \mid B_j\right) P\left(B_j\right)$ e então teremos o que se denomina o **teorema de probabilidade total**:
$$P(A) = P\left(A \mid B_1\right) P\left(B_1\right) + P\left(A \mid B_2\right) P\left(B_2\right) + \cdots + P\left(A \mid B_k\right) P\left(B_k\right)$$

Teorema de Bayes
Motivação

- Uma determinada peça é manufaturada por **três fábricas**, digamos 1, 2 e 3.
- Sabe-se que 1 produz o **dobro** de peças que 2, e 2 e 3 produziram **o mesmo número** de peças (durante um período de produção especificado).
- Sabe-se também que **2%** das peças produzidas por 1 e por 2 são defeituosas, enquanto **4%** daquelas produzidas por 3 são defeituosas.
- Todas as peças produzidas são colocadas em um depósito, e depois uma peça é extraída **ao acaso**.
- Vamos introduzir os seguintes eventos:

$$A = \{\text{a peça é defeituosa}\},\, B_1 = \{\text{a peça provém de } 1\},\; B_2 = \{\text{a peça provém de } 2\},$$
$$B_3 = \{\text{a peça provém de } 3\}$$

- Qual é a probabilidade de que essa **peça seja defeituosa**?

Teorema de Bayes
Motivação

- Pede-se $P(A)$, e **empregando-se o teorema da probabilidade total** poderemos escrever:

$$P(A) = P(A \mid B_1) P(B_1) + P(A \mid B_2) P(B_2) + P(A \mid B_3) P(B_3)$$

- Temos $P(B_1) = 1/2$, enquanto $P(B_2) = P(B_3) = 1/4$.
- Além disso, $P(A \mid B_1) = P(A \mid B_2) = 0{,}02$, enquanto $P(A \mid B_3) = 0{,}04$.
- Substituindo estes valores na expressão acima, temos $P(A) = 0{,}025$.
- Agora considere um novo problema: suponha-se que uma peça seja retirada do depósito e verifica-se que ela é defeituosa. **Qual é a probabilidade de que tenha sido produzida na fábrica 1**?
- Empregando a notação já introduzida, pede-se $P(B_1 \mid A)$.

Teorema de Bayes
Motivação

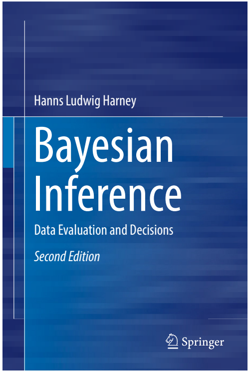
- Seja $B_1, B_2, \ldots, B_k$ uma **partição** do espaço amostral S e seja A um **evento** associado a S.
- Aplicando-se a definição de probabilidade condicionada, podemos escrever

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{\sum_{j=1}^{k} P\left(A \mid B_j\right) P\left(B_j\right)}, \quad i = 1, 2, \ldots, k$$

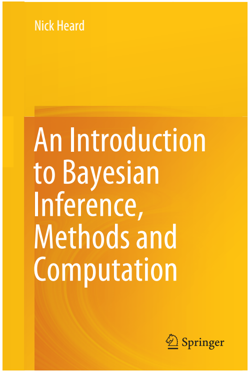
- Este resultado é conhecido como **Teorema de Bayes**.
- Desde que os B_i constituam uma partição do espaço amostral **um, e somente um**, dos eventos B_i ocorrerá.
- Portanto, a expressão acima nos dá a probabilidade de um B_i **particular** (isto é, uma "causa"), **dado que** o evento A tenha ocorrido.
- A fim de aplicar esse teorema, deveremos **conhecer** os valores das $P(B_i)$.
- Para o problema proposto, $P(B_i \mid A) = 0{,}40$. **Verifique este resultado.**
- **Na próxima aula, vamos falar sobre variáveis aleatórias e medidas de dispersão.**

Teorema de Bayes

Sugestões para consultar uma apresentação mais formal



[1] H. L. Harney.
Bayesian Inference: Data Evaluation and Decisions.
Springer International Publishing, Cham, 2 edition, 2016.



[1] N. Heard.
An Introduction to Bayesian Inference, Methods and Computation.
Springer, Cham, 1 edition, 2021.