1^a Prova - FECD A - 2025

Renato Assunção

1. Uma variável aleatória discreta X assume os valores $1, 2, 3, 4$ com probabilidades

$$\mathbb{P}(X = x) = k \cdot x, \quad x = 1, 2, 3, 4.$$

- Determine a constante de normalização k.
- Obtenha a função de distribuição acumulada $F(x) = \mathbb{P}(X \leq x)$.
- Calcule $\mathbb{E}(X)$.

Solução: $1 = \sum_{x=1}^{4} kx = k(1 + 2 + 3 + 4) = 10k \Rightarrow k = 0.1$.
A função de distribuição acumulada $F(x) = \mathbb{P}(X \leq x)$ para todo $x \in \mathbb{R}$ é:

$$F(x) = \begin{cases} 0, & x < 1, \\ 0{,}1, & 1 \leq x < 2, \\ 0{,}3, & 2 \leq x < 3, \\ 0{,}6, & 3 \leq x < 4, \\ 1, & x \geq 4. \end{cases}$$

Observação: F salta nos pontos $1, 2, 3, 4$ e os saltos têm tamanhos iguais a $\mathbb{P}(X = x) = 0{,}1x$.
$\mathbb{E}(X) = \sum x \cdot kx = 0.1(1 + 4 + 9 + 16) = 3$.

2. Um pesquisador está interessado em estudar a relação causal entre *tempo de tela diário* (horas em celular/computador) e *qualidade do sono* entre estudantes. Ele realiza um estudo observacional por questionários e encontra que alunos com *menos* tempo de tela têm, em média, melhor sono autorrelatado. Explique por que é difícil concluir, a partir dessa evidência, que reduzir o tempo de tela *causa* melhora do sono. Dê dois exemplos de possíveis confundidores e explique o papel do desenho amostral nessa dificuldade.

Solução (ideias-chave): Fatores de confundimento: rotinas familiares/horário de trabalho dos pais; estresse/saúde mental; disciplina de estudos; condição socioeconômica; atividade física; café/estimulantes. Viés de seleção/medida. Necessidade de desenho experimental (randomização) ou controles/estratificação/modelagem.

3. Sejam A e B eventos com $\mathbb{P}(B) > 0$. Mostre que

- $\mathbb{P}(A^c \mid B) = 1 - \mathbb{P}(A \mid B)$
- $\mathbb{P}(A \cup C \mid B) = \mathbb{P}(A \mid B) + \mathbb{P}(C \mid B)$ se A e C são disjuntos

Solução: Use a definição de probabilidade condicional e o fato de que A e A^c são disjuntos e $A \cup A^c = \Omega$.

$$\mathbb{P}(A \mid B) + \mathbb{P}(A^c \mid B) = \frac{\mathbb{P}((A \cap B) \cup (A^c \cap B))}{\mathbb{P}(B)} = \frac{\mathbb{P}((A \cup A^c) \cap B))}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$

$$\begin{aligned} \mathbb{P}(A \cup C \mid B) &= \frac{\mathbb{P}((A \cup C) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}((A \cap B) \cup (C \cap B))}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(A \cap B) + \mathbb{P}(C \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A \mid B) + \mathbb{P}(C \mid B) \end{aligned}$$

4. Em um estudo sobre o desempenho de alunos em Matemática, eles foram classificados de acordo com o **tipo de escola** (4 tipos: Pública Estadual, Pública Municipal, Privada, Técnica) e o **nível de desempenho** (3 categorias: Alto, Médio, Baixo). As probabilidades de que um aluno caia em cada uma das 12 categorias são dadas na tabela abaixo:

Escola	Alto	Médio	Baixo
Pública Estadual	0.12	0.08	0.10
Pública Municipal	0.10	0.07	0.08
Privada	0.20	0.10	0.05
Técnica	0.07	0.02	0.01

Calcule as seguintes probabilidades (pode deixar as contas apenas indicadas):

(a) $\mathbb{P}(\text{Médio})$

(b) $\mathbb{P}(\text{Privada})$

(c) $\mathbb{P}(\text{Privada} \mid \text{Médio})$

(d) $\mathbb{P}(\text{Médio} \mid \text{Privada})$

Solução:

1. $\mathbb{P}(\text{Médio}) = 0.08 + 0.07 + 0.10 + 0.03 = 0.28$
2. $\mathbb{P}(\text{Privada}) = 0.20 + 0.10 + 0.05 = 0.35$
3. $\mathbb{P}(\text{Privada} \mid \text{Médio}) = \dfrac{0.10}{0.28} \approx 0.357$
4. $\mathbb{P}(\text{Médio} \mid \text{Privada}) = \dfrac{0.10}{0.35} \approx 0.286$

5. Em uma população, 2% das pessoas possuem uma certa doença rara. Existe um teste médico para detectar essa doença. O teste é positivo em 95% dos doentes (sensibilidade) e também positivo em 5% dos saudáveis (falsos positivos).

   (a) Se uma pessoa sorteada ao acaso tem o teste positivo, qual a probabilidade de ela realmente ter a doença?

   (b) Ao fazer o cálculo, você vai obter a resposta igual a 0.28, Interprete este resultado e explique a importância da prevalência na aplicação da regra de Bayes.

Solução: Como $\mathbb{P}(\text{doente}) = 0.02$, temo $\mathbb{P}(\text{sadio}) = 1 - \mathbb{P}(\text{doente}) = 0.98$. O problema fornece as probabilidades:

$$\mathbb{P}(+ \mid \text{doente}) = 0.95, \qquad \mathbb{P}(+ \mid \text{sadio}) = 0.05\,.$$

Pela regra de Bayes,

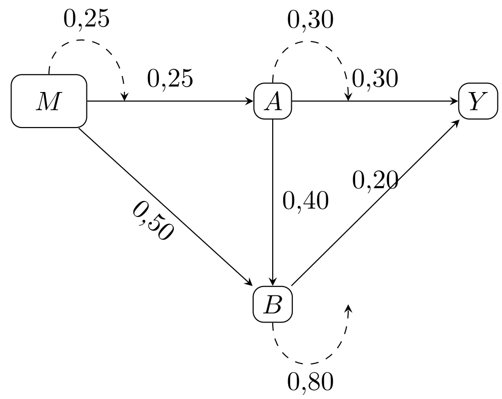
$$\begin{aligned}
\mathbb{P}(\text{doente} \mid +) &= \frac{\mathbb{P}(+ \mid \text{doente})\mathbb{P}(\text{doente})}{\mathbb{P}(+)} \\
&= \frac{0.95 \cdot 0.02}{\mathbb{P}(+ \mid \text{doente})\mathbb{P}(\text{doente}) + \mathbb{P}(+ \mid \text{sadio})\mathbb{P}(\text{sadio})} \\
&= 0.019/(0.95 \cdot 0.02 + 0.05 \cdot 0.98) = 0.019/0.068 \approx 0.28
\end{aligned}$$

O resultado mostra que, mesmo com um teste muito sensível (95%) e relativamente específico (95%), a probabilidade de uma pessoa estar realmente doente dado que o teste foi positivo é de apenas cerca de 28%. Isso ocorre porque a doença é muito rara (prevalência de apenas 2%).

Quando a prevalência é baixa, o número de falsos positivos (indivíduos saudáveis cujo teste dá positivo) pode ser maior do que o número de verdadeiros positivos, fazendo com que a probabilidade pós-teste seja relativamente pequena. Esse exemplo ilustra que o valor preditivo de um teste não depende apenas da sensibilidade e da especificidade do exame, mas também da prevalência da condição na população.

6. Um programa possui quatro módulos: M, A, B e Y. A entrada ocorre apenas em M. Cada módulo pode ser executado *no máximo uma vez* e o programa pode ser interrompido em qualquer módulo. As transições (com probabilidades condicionais) são:

   - De M: vai para A com prob. 0,25, para B com prob. 0,50, ou interrompe com prob. 0,25.
   - De A: vai para Y com prob. 0,30, para B com prob. 0,40, ou interrompe com prob. 0,30.
   - De B: vai para Y com prob. 0,20 ou interrompe com prob. 0,80.
   - Y é terminal.



(a) Calcule $\mathbb{P}(Y \mid A \text{ foi executado})$.

(b) Calcule $\mathbb{P}(A \mid B \text{ foi executado})$.

Solução: (a) Estando em A, caminhos até Y: $A \to Y$ (prob. 0,30) ou $A \to B \to Y$ (prob. $0{,}40 \times 0{,}20 = 0{,}08$). Como esses caminhos representam eventos disjuntos $\Rightarrow \mathbb{P}(Y \mid A) = 0{,}30 + 0{,}08 = 0{,}38$.

(b) Temos $\mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. Cada módulo ocorre no máximo uma vez e as maneiras de chegar a B são $M \to B$ e $M \to A \to B$. Apenas o segundo tem A e B. Assim,

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(M \to A \to B)}{\mathbb{P}(B)}$$

Temos $\mathbb{P}(M \to A \to B) = 0{,}25 \times 0{,}40 = 0{,}10$.

Temos também $\mathbb{P}(B) = \mathbb{P}(M \to B) + \mathbb{P}(M \to A \to B) = 0{,}50 + 0{,}10 = 0{,}60$.

Assim, $\mathbb{P}(A \mid B) = 0{,}10/0{,}60 = \frac{1}{6} \approx 0{,}167$.