Universidade de São Paulo
Instituto de Matemática e Estatística
Aluna: Nara Avila Moraes
Nro. USP: 5716734

2o. semestre, 2025
Exercício 01 - 08/09

MAC5911/IME: Fundamentos de Estatística e Machine Learning. Prof.: Alexandre Galvão Patriota

Questão 01: Apresente um texto de no máximo duas páginas que introduza uma medida de possibilidade condicional, incluindo pelo menos um exemplo numérico e um teorema. Sugiro que leia o paper do Friedman e Halpern (1995) e busque referências adicionais sobre o assunto que estejam publicadas em revistas internacionais. Por exemplo, os autores Didier Dubois e Henry Prade estudaram o assunto em vários artigos.

Questão 02: Na discussão sobre 'The Dutch Book Argument', considere um jogador que não utiliza probabilidades e a banca escolhe uma configuração para explorar a perda certa que o jogador terá. Apresente:

(2.1) Os valores numéricos de P(H,E) diferentes dos discutidos em sala para cada "H1", "H2", e "H1 U H2";

(2.2) As apostas escolhidas pela banca para explorar a perda certa do jogador;

(2.3) A tabela demonstrando que, em todas as possibilidades, o jogador perde para a banca;

(2.4) Comentários sobre os resultados.

Resposta para o Questão 01

Introdução à Medida de Possibilidade Condicional

Embora a probabilidade seja a medida clássica da incerteza sobre a realidade física, ela não é suficiente para caracterizar a forma de raciocínio epistêmico humano, também chamado de raciocínio não monotônico. Desde Aristóteles, passando por economistas do século XIX, filósofos e engenheiros modernos, observou-se que muitos sistemas — especialmente sistemas de inteligência artificial e o raciocínio humano — não estão contidos no arcabouço probabilístico. Entre as diferenças fundamentais, destacam-se a assimetria, a não monotonicidade e a não associatividade como regra, presentes na racionalização. Nesses casos, a incerteza não provém da aleatoriedade, mas sim da incompletude ou da imprecisão da informação. Fisher foi o primeiro a buscar uma caracterização axiomática para o que ele cunhou de o termo mais geral possível, dentre todos já definidos anteriormente: a **medida de plausibilidade**. Trata-se de um conceito mais abrangente, do qual a probabilidade e a possibilidade emergem como casos particulares.

Algebricamente, Fisher define uma medida de plausibilidade como a tupla

$$(\mathcal{W}, \mathcal{F}, \mathcal{D}, \mathcal{P}l),$$

em que:

- $\mathcal{W}$ é o conjunto de mundos possíveis;
- $\mathcal{F} = 2^{\mathcal{W}}$ é a família de subconjuntos de $\mathcal{W}$ (o conjunto dos eventos);
- $\mathcal{D}$ é um conjunto parcialmente ordenado, que permite definir extensão, transitividade e comparação ordinal entre graus de plausibilidade;
- $\mathcal{P}l : \mathcal{F} \to \mathcal{D}$ é a função que atribui a cada evento um grau de plausibilidade.

A diferença central para a probabilidade é o tipo de decomposição permitido:

- Para a probabilidade, a decomposição de eventos mutuamente exclusivos se dá pela soma:

$$P(A \cup B) = P(A) + P(B), \quad \text{se } A \cap B = \emptyset.$$

- Para a possibilidade, a decomposição ocorre pela máxima:

$$\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}.$$

Essa diferença gera uma **assimetria estrutural**: a medida de possibilidade é decomposta pela união, mas sua dual — a medida de necessidade — só se decompõe pela interseção:

$$N(A \cap B) = \min\{N(A), N(B)\}.$$

Já a probabilidade é simétrica: é ao mesmo tempo sua própria dual e se decompõe tanto pela união quanto pela interseção, obedecendo às regras clássicas de aditividade. Assim surgiu a Teoria da Possibilidade, introduzida por L. A. Zadeh, para lidar com a incerteza epistêmica. A teoria baseia-se em uma distribuição de possibilidade, $\pi : \Omega \to [0, 1]$, que atribui a cada elemento ω do universo de discurso Ω um grau de possibilidade, onde $\pi(\omega) = 1$ significa que ω é totalmente possível e $\pi(\omega) = 0$ significa que é impossível. A partir de π, duas medidas duais são definidas para qualquer evento $A \subseteq \Omega$:

- **Medida de Possibilidade (Π)**: Avalia o grau em que o evento A é consistente com a informação disponível. É definida como $\Pi(A) = \sup_{\omega \in A} \pi(\omega)$. Esta medida satisfaz a propriedade axiomática: $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$.
- **Medida de Necessidade (N)**: Avalia o grau em que o evento A é certamente implicado pela informação. É definida como $N(A) = 1 - \Pi(A^c)$, onde A^c é o complementar de A.

A grande vantagem deste formalismo é a sua capacidade de distinguir entre a falta de crença e a descrença. Se $N(A) = 0$, não significa que A é falso, mas apenas que não há evidência que o torne necessário.

Condicionamento Possibilístico: Atualizando Crenças

Assim como a probabilidade condicional é essencial para a atualização de crenças no modelo probabilístico, a possibilidade condicional é crucial para a revisão de crenças possibilísticas quando uma nova informação, um evento B, é observada. O objetivo é definir $\Pi(A|B)$, o grau de possibilidade de um evento A dado que B ocorreu.

A relação fundamental é:

$$\Pi(A \cap B) = \min(\Pi(A|B), \Pi(B)).$$

Definição: A medida de possibilidade condicional de um evento A dado um evento B, com $\Pi(B) > 0$, é definida por:

$$\Pi(A|B) = \begin{cases} 1 & \text{se } \Pi(A \cap B) = \Pi(B), \\ \Pi(A \cap B) & \text{se } \Pi(A \cap B) < \Pi(B). \end{cases}$$

Como principal teorema podemos citar a Lei da Possibilidade Total. De forma análoga à Lei da Probabilidade Total, existe um teorema correspondente na Teoria da Possibilidade:

Teorema (Lei da Possibilidade Total): Seja $\{B_1, B_2, \ldots, B_n\}$ uma partição do universo Ω. Então, para qualquer evento $A \subseteq \Omega$, sua possibilidade incondicional pode ser calculada a partir das possibilidades condicionais:

$$\Pi(A) = \max_{i=1,\ldots,n} \Pi(A \cap B_i) = \max_{i=1,\ldots,n} \min(\Pi(A|B_i), \Pi(B_i)).$$

Diagnóstico Médico

Para exemplificar, suponha que um paciente pode ter uma de três doenças mutuamente exclusivas: D_1 (Gripe), D_2 (Virose Comum) ou D_3 (Alergia). A distribuição inicial é:

$$\pi(D_1) = 1.0, \quad \pi(D_2) = 0.8, \quad \pi(D_3) = 0.4.$$

Um sintoma S (febre alta) é observado, com possibilidades condicionais:

$$\Pi(S|D_1) = 0.9, \quad \Pi(S|D_2) = 0.5, \quad \Pi(S|D_3) = 0.1.$$

A possibilidade do sintoma S:

$$\Pi(S) = \max\{\min(0.9, 1.0), \min(0.5, 0.8), \min(0.1, 0.4)\} = \max\{0.9, 0.5, 0.1\} = 0.9.$$

As possibilidades atualizadas (posteriores):

- Para D_1: $\Pi(D_1|S) = 1.0$.
- Para D_2: $\Pi(D_2|S) = 0.5$.
- Para D_3: $\Pi(D_3|S) = 0.1$.

Após observar febre alta, a nova distribuição é:

$$\pi(D_1|S) = 1.0, \quad \pi(D_2|S) = 0.5, \quad \pi(D_3|S) = 0.1.$$

Resposta para o Ítem 2.1

Consideremos dois eventos mutuamente exclusivos, H_1 e H_2. Um jogador, cujas atribuições de probabilidade não são coerentes, define os seguintes valores, diferentes de uma atribuição clássica que usaria a frequência:

- $P(H_1) = 0.4$
- $P(H_2) = 0.5$
- $P(H_1 \cup H_2) = 0.8$

A incoerência aqui reside no fato de que, para eventos mutuamente exclusivos, a soma das probabilidades deve ser

$$P(H_1 \cup H_2) = P(H_1) + P(H_2).$$

No nosso caso,

$$0.4 + 0.5 = 0.9,$$

que é diferente de 0.8.

Resposta para o Ítem 2.2

Para explorar a incoerência do jogador, a banca oferece as seguintes apostas, assumindo que cada aposta tem valor nominal de R$ 1,00:

1. **Aposta 1:** A banca compra do jogador uma aposta sobre o evento H_1, ao preço (probabilidade) de $P(H_1) = 0.4$. A aposta paga R$ 1,00 se H_1 não ocorrer e o jogador perde R$ 1,00 se H_1 ocorrer.
2. **Aposta 2:** A banca compra do jogador uma aposta sobre o evento H_2, ao preço (probabilidade) de $P(H_2) = 0.5$. A aposta paga R$ 1,00 se H_2 não ocorrer e o jogador perde R$ 1,00 se H_2 ocorrer.
3. **Aposta 3:** A banca vende ao jogador uma aposta sobre o evento $H_1 \cup H_2$, ao preço (probabilidade) de $P(H_1 \cup H_2) = 0.8$. A aposta paga R$ 1,00 se $H_1 \cup H_2$ ocorrer e o jogador perde R$ 1,00 se não ocorrer.

Em outras palavras, o jogador está disposto a:

- Pagar R$ 0,40 para ter a chance de ganhar R$ 1,00 se H_1 ocorrer.
- Pagar R$ 0,50 para ter a chance de ganhar R$ 1,00 se H_2 ocorrer.

- Vender uma aposta em $H_1 \cup H_2$ por R$ 0,80.

O lucro inicial do jogador ao aceitar todas as apostas é:

$$\text{Lucro Inicial} = (\text{Pagamento pela Aposta 3}) - (\text{Custo da Aposta 1}) - (\text{Custo da Aposta 2})$$

$$\text{Lucro Inicial} = 0.80 - 0.40 - 0.50 = -0.10$$

O jogador já começa a transação com um prejuízo de R$ 0,10.

Resposta para o Ítem 2.3

A tabela a seguir demonstra o resultado final do lucro ou prejuízo do jogador, considerando todos os cenários possíveis para os eventos H_1 e H_2, que são mutuamente exclusivos:

Resultado	**Lucro (Aposta 1)**	**Lucro (Aposta 2)**	**Lucro (Aposta 3)**	**Lucro Total**
H_1 ocorre	$+0.60$	-0.50	-0.20	-0.10
H_2 ocorre	-0.40	$+0.50$	-0.20	-0.10
Nenhum ocorre	-0.40	-0.50	$+0.80$	-0.10

A tabela confirma que, independentemente do resultado dos eventos, o jogador sempre terá um prejuízo líquido de R$ 0,10.

Resposta para o Ítem 2.4

O *"Dutch Book Argument"* evidencia a importância da **coerência formal** na atribuição de probabilidades. Do ponto de vista de um estatístico clássico, o jogador não perde por causa de uma crença subjetiva, mas porque os valores que ele escolheu para suas "probabilidades" não formam uma medida de probabilidade válida.

A teoria da probabilidade, construída sobre axiomas como a aditividade finita, é uma estrutura matemática rigorosa. Qualquer sistema de atribuição de valores que viole esses axiomas é, por definição, internamente inconsistente. A perda certa, neste caso, não é uma punição, mas a consequência lógica e inevitável de operar fora das regras da lógica probabilística. O argumento reforça o princípio de que a probabilidade, para ser útil e não levar a contradições, deve ser tratada como uma medida objetiva e não como uma simples manifestação de crença pessoal.