21073 - Introdução às probabilidades e estatística bayesianas

Ano lectivo 2016/16

Docente: António Araújo

e-fólio A

Para a resolução do e-fólio, aconselha-se que:

- Verifique se o ficheiro que recebeu está correcto. O e-fólio consiste de 2 páginas com 5 problemas e termina com a palavra FIM.
- Como o e-fólio tem um tempo prolongado de resolução, espera-se que as respostas que enviar estejam legíveis, com boa apresentação e organização. Deve fazer à parte o trabalho auxiliar e enviar apenas uma versão final, "limpa". Deve digitalizar a sua resolução de forma legível, ou executá-la directamente em formato digital (aceita-se word, pdf, ou scans em jpeg, png ou tiff - se usar varios ficheiros envie apenas um arquivo com todos eles, em rar ou zip). Respostas ilegíveis não serão cotadas, por isso verifique bem o seu ficheiro antes de enviar.
- Justifique cuidadosamente todas as suas respostas. Apresente todos os cálculos que julgue necessários para a compreensão do seu raciocínio.
- Tenha em atenção o prazo de entrega do e-fólio e as indicações para submeter a resolução disponibilizadas na sala de aulas virtual.
- O e-fólio é um trabalho individual. Pode utilizar recursos externos (pesquisa online, literatura, etc) mas não pode pedir ajuda a terceiros nem discutir os problemas com os seus colegas.

Critérios de avaliação e cotação:

- Este e-fólio tem a cotação total de 4 valores, assim distribuídos: cada questão vale 1 valor.

Por favor preencha os seus dados:

- Nome:
- B.I:
- Nº de Estudante
- Curso:

1. Considere a seguinte proposta de silogismo fraco:

$$p(B|\overline{A}, A \Rightarrow B) \leq p(B|A \Rightarrow B)$$

Descreva o seu significado "por palavras". Será verdadeiro? Demonstre que sim ou que não.

Solução:

O silogismo fraco em causa é

$$\frac{\begin{array}{c}\text{A implica B}\\ \text{A é falso}\end{array}}{\text{Então, B é menos provável}}$$

Ou, mais por extenso: "Se sabemos que 'A implica B' é verdade e se sabemos que A é falso então B torna-se menos provável do que se apenas sabemos que 'A implica B' é verdade".

Vamos mostrar que este silogismo é verdadeiro:

$$p(B|\overline{A}, A \Rightarrow B) = \frac{p(\overline{A}|B, A \Rightarrow B)p(B|A \Rightarrow B)}{p(\overline{A}|A \Rightarrow B)} < p(B|A \Rightarrow B)$$

onde o último passo decorre da validade do silogismo fraco

$p(A|A \Rightarrow B, B) \geq p(A|A \Rightarrow B)$. De facto, de $p(A|A \Rightarrow B, B) \geq p(A|A \Rightarrow B)$ tiramos que $1 - p(A|A \Rightarrow B, B) \leq 1 - p(A|A \Rightarrow B)$, e portanto $p(\overline{A}|A \Rightarrow B, B) \leq p(\overline{A}|A \Rightarrow B)$, ou seja

$$\frac{p(\overline{A}|A \Rightarrow B, B)}{p(\overline{A}|A \Rightarrow B)} \leq 1.$$

2. Suponha que A_1, A_2, X, C são proposições, e que

1) $p(A_1A_2|CX) = p(A_1|CX)p(A_2|CX)$,
2) $p(C|X) = 0.5,\ p(A_1|CX) = 0.1,\ p(A_2|CX) = 0.2$
3) $p(A_1|\overline{C}X) = 0.5,\ p(A_2|\overline{C}X) = 0.4$.
Calcule $p(C|A_1A_2X)$.

Solução:

$$p(C|A_1A_2X) = \frac{p(A_1A_2|CX)p(C|X)}{p(A_1A_2|X)}$$
$$= \frac{p(A_1|CX)p(A_2|CX)p(C|X)}{p(A_1A_2|X))} = \frac{p(A_1|CX)p(A_2|CX)p(C|X)}{p(A_1A_2(C+\overline{C})|X)}$$
$$= \frac{p(A_1|CX)p(A_2|CX)p(C|X)}{p(A_1A_2C|X) + p(A_1A_2\overline{C}|X)} = \frac{p(A_1|CX)p(A_2|CX)p(C|X)}{p(A_1A_2|CX)p(C|X) + p(A_1A_2|\overline{C}X)p(\overline{C}|X)}$$
$$= \frac{p(A_1|CX)p(A_2|CX)p(C|X)}{p(A_1|CX)p(A_2|CX)p(C|X) + p(A_1|\overline{C}X)p(A_2|\overline{C}X)p(\overline{C}|X)}$$

Substituindo os valores obtemos

$$p(C|A_1A_2X) = \frac{0.1 \cdot 0.2 \cdot 0.5}{0.1 \cdot 0.2 \cdot 0.5 + 0.5 \cdot 0.4 \cdot (1 - 0.5)} = \frac{10}{110} \cong 0.09$$

3. Sejam $A_1, \ldots, A_n, X, C$ proposições. Suponhamos que, dado X, as A_i são mutuamente exclusivas, isto é, $p(A_iA_j|X) = 0$ para todo o $i \neq j$.

Mostre que

$$p(C|(A_1 + \ldots + A_n)X) = \frac{\sum_i p(A_i|X)p(C|A_iX)}{\sum_i p(A_i|X)}.$$

Solução:

$$p(C|(A_1 + \ldots + A_n)X) = p((A_1 + \ldots + A_n)|CX).\frac{p(C|X)}{p((A_1 + \ldots + A_n)|X)}$$
$$= \frac{\Big(p(A_1|CX) + \ldots + p(A_n|CX)\Big).p(C|X)}{p(A_1 + \ldots + A_n|X)}$$
$$= \frac{p(A_1|X)p(C|A_1X) + \ldots + p(A_n|X).p(C|A_nX)}{p(A_1|X) + \ldots + p(A_n|X)}$$
$$= p(C|(A_1 + \ldots + A_n)X) = \frac{\sum_i p(A_i|X)p(C|A_iX)}{\sum_i p(A_i|X)}$$

onde usámos, no terceiro passo, o facto de que

$$P(A_i|CX) = \frac{p(A_i|X).p(C|A_iX)}{p(C|X)}$$

para eliminar $p(C|X)$.

4. Sejam X, Y, Z proposições. Diz-se que X é independente de Y sabendo (ou "condicionado a") Z se $p(X|YZ) = p(X|Z)$. Mostre que

a) X é independente de Y sabendo Z se e só se $p(XY|Z) = p(X|Z)(Y|Z)$.

b) Se X é independente de Y sabendo Z então Y é independente de X sabendo Z.

c) Se X é independente de Y sabendo Z então não-X também é independente de Y sabendo Z.

Solução:

a) Supondo independência, temos $p(XY|Z) = p(X|YZ)p(Y|Z) = p(X|Z)p(Y|Z)$. Recíprocamente, assumindo que $p(XY|Z) = p(X|Z)p(Y|Z)$, vem que $p(X|YZ) = p(XY|Z)/p(Y|Z) = p(X|Z)p(Y|Z)/p(Y|Z) = p(X|Z)$.

b) Feita a alínea a) torna-se um mero corolário da simetria do produto.

c)

$$\begin{aligned}
&p(\overline{X}Y|Z) = p(\overline{X}|YZ)p(Y|Z) = (1 - p(X|YZ))p(Y|Z) \\
=&p(Y|Z) - p(X|YZ)p(Y|Z) \overset{\text{(hip.)}}{=} p(Y|Z) - p(X|Z)p(Y|Z) \\
=&p(Y|Z)(1 - p(X|Z)) = p(Y|Z)p(\overline{X}|Z)
\end{aligned}$$

FIM