

Teoria da Informação

Ficha Teórico-Prática nº 1

“Teoria da Informação”

Período de execução: 1 aula teórico-prática

Objectivo: Pretende-se que o aluno adquira sensibilidade para as questões fundamentais de teoria da informação, em particular informação, entropia exacta e aproximada e incerteza.

Trabalho

1. Numa imagem binária 20x20, 200 pixels são pretos.
 - a) Sem efectuar quaisquer cálculos, indique a informação e a entropia de cada pixel.
 - b) Considere agora que 240 pixels são pretos. Calcule a informação e a entropia de cada pixel (em bits). Qual a relação entre a informação de cada pixel e a probabilidade da sua ocorrência?
 - c) Nessa mesma imagem, verificou que a probabilidade de ocorrência de 2 pixies consecutivos <preto, preto> é de 0.6, <branco, branco> é de 0.2, <preto, branco> é de 0.1 e <branco, preto> é de 0.1. Calcule a entropia de cada pixel. Compare-a com a calculada na alínea anterior e tire conclusões.
 - d) Assumindo agora que 2/3 dos pixeis são pretos, verificou que a probabilidade de um pixel preto suceder a outro pixel preto é 0.6, um branco suceder a um branco é de 0.2 e um preto suceder a um branco é de 0.8 e um branco suceder a um preto é de 0.4. Calcule a entropia associada a cada pixel.
2. Considere uma fonte de informação com a seguinte estatística de segunda ordem ($P(x,y)$)

x/y	0	1
0	1/3	1/3
1	0	1/3

- a) Determine $H(X)$ e $H(Y)$.
 - b) Determine $H(X|Y)$ e $H(Y|X)$.
 - c) Determine $H(X,Y)$
 - d) Determine $I(X;Y)$
3. Demonstre as seguintes propriedades elementares:
 - a) $H(x) \geq 0$
 - b) $H(x) \leq \log_2(N)$, em que N representa a cardinalidade do alfabeto da fonte de informação.

c) $H(x,y) = H(x) + H(y)$ se e só se x e y forem variáveis independentes.

4. Considere a seguinte fonte de informação, X , que regista as temperaturas ao longo do dia com intervalo de duas horas:

12 14 16 18 20 22 22 20 18 16 14 12

- a) Calcule a entropia de fonte.
- b) Proponha uma estratégia de modelação da fonte de informação que minimize o valor de entropia.
- c) a) Para aprender o processo que esteve na origem de X pretende-se que a saída Y do algoritmo de aprendizagem verifique:
 - ☐ $H(Y) > H(X)$
 - ☐ $I(X;Y) = 0$
 - ☐ $I(X;Y) = H(X)$
 - ☐ nenhuma das anteriores

5. Considere uma fonte de informação composta por um alfabeto de dois símbolos $S=\{1,2\}$. Assuma que as probabilidades condicionadas de ordem 1 dessa fonte são $P(1|1)=0.8$, $P(2|1)=0.2$, $P(1|2)=0.6$ e $P(2|2)=0.4$. (note que $P(1,2)=P(2,1)$)

- a) Comente a seguinte afirmação: "Um codificador que assuma que os símbolos são independentes terá que usar, em média, pelo menos 0.95 bits por símbolo." Fundamente com cálculos.
- b) Considere que procede à modelização da sua fonte de informação usando um modelo de Markov de primeira ordem. Nessas circunstâncias, qual é o melhor desempenho possível do codificador.

6. Considere uma variável estocástica $X \in \{1,2,3\}$ tal que $P(X=1)=a/2$, $P(X=2)=1-a$ e $P(X=3)=a/2$, $a \in [0,1]$. Nestas circunstâncias observa-se que (assinale a(s) resposta(s) certa(s)):

- a) $H(X)$ é máximo para:
 - ☐ $a = 1$
 - ☐ $a = 1/3$
 - ☐ $a = 1/2$
 - ☐ nenhuma das anteriores
- b)
 - ☐ $H(X) = 1$
 - ☐ $H(X) \leq 2$
 - ☐ $H(X) \geq 0$
 - ☐ nenhuma das anteriores
- c)
 - ☐ $H(X,X) = H(X)$
 - ☐ $H(X,X) = H(X) + H(X)$
 - ☐ $H(X,X) = H(X) + H(X|X)$
 - ☐ nenhuma das anteriores
- c)
 - ☐ $I(X;X) = 0$
 - ☐ $I(X;X) = H(X|X)$
 - ☐ $I(X;X) = H(X)$
 - ☐ nenhuma das anteriores
- d)
 - ☐ $I(X;X)$ é máximo quando $a = 0$
 - ☐ $I(X;X)$ é máximo quando $a = 1$
 - ☐ $I(X;X)$ é máximo quando $a = 0.5$
 - ☐ nenhuma das anteriores

7. Considere que $Y = \log_2(2X+2)$. Assumindo que $D()$ representa a distância KL, assinale as opções corretas:

- a) ☐ $H(X,Y) \geq 0$ ☐ $H(X,Y) \leq 1$
☐ $H(Y) \geq 0$ ☐ nenhuma das anteriores
- b) ☐ $H(X,Y) = H(Y)$ ☐ $H(X,Y) = H(Y) + H(X)$
☐ $H(X,Y) = H(X) + H(Y|X)$ ☐ nenhuma das anteriores
- c) ☐ $I(X;Y) = 0$ ☐ $I(X;Y) = H(X|Y)$
☐ $I(X;Y) = H(Y)$ ☐ nenhuma das anteriores
- d) ☐ $D(X;Y) = 0$ ☐ $D(X;Y) + D(Y;X) \leq 2$
☐ $D(X;Y) + D(Y;X) = H(Y) + H(X)$ ☐ nenhuma das anteriores
- e) ☐ $D(X;Y) + D(Y;X) = 2H(X)$ ☐ $D(X;Y) + D(Y;X) \geq 0$
☐ $D(X;Y) + D(Y;X) = 2H(Y)$ ☐ nenhuma das anteriores

8. Considere um esquema de codificação de canal em que o codificador triplica cada bit transmitido, isto é, cada bit t o codificador codifica $t1t2t3$. O decodificador para cada sequência $r1r2r3$ decodifica a mensagem $s = \text{mediana}\{r1r2r3\}$. Assuma que a probabilidade de erro de transmissão de cada bit é de 10% e que os acontecimentos são independentes. Determine a probabilidade de erro de cada mensagem.

9*. Uma moeda de 1€ é lançada até que ocorra a primeira cara.

- a) Sendo X o número requerido de lançamentos, calcule a entropia de X em bits. Assuma a situação genérica em que a moeda possa estar viciada, sendo f a probabilidade de ocorrência de caras.
- b) Qual a entropia quando a moeda é totalmente equilibrada?

Utilize os seguintes resultados:

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r} \qquad \sum_{n=1}^{\infty} nr^n = \frac{r}{(1-r)^2}$$

10**. A final da NBA consiste numa série de 7 jogos que termina logo que uma das equipas vença 4 jogos. Seja X uma variável aleatória que representa o resultado de uma final entre as equipas A e B; exemplos de possíveis valores de X são: AAAA, BABABB, BBBAAAA, etc. Seja Y o número de jogos realizados, que varia entre 4 e 7. Assumindo que as equipas A e B são equilibradas e que os jogos são independentes, calcule a entropia de X , $H(X)$, a entropia de Y , $H(Y)$, a entropia condicional conhecendo-se Y , $H(X|Y)$, a entropia condicional conhecendo-se X , $H(Y|X)$ e a informação mútua, $I(X, Y)$. Tire conclusões.

* Exercício adaptado de Cover and Thomas, "Elements of Information Theory", p. 42.

** Exercício adaptado de Cover and Thomas, "Elements of Information Theory", p. 44.

11. Demonstre, com recurso ao princípio da máxima entropia, que a entropia de uma dada variável aleatória X é máxima quando os acontecimentos são equiprováveis.