Departamento de Engenharia Informática
Faculdade de Ciências e Tecnologia
Universidade de Coimbra

Teoria da Informação

16, Janeiro 2023

Exame Época Recurso

Duração: 2h

NOME: Nº Estudante:

Observe que:

Exame com consulta condicionada (uma página A4 de apontamentos)
Não são permitidos meios electrónicos (computador, telemóveis, etc.), excepto calculadoras.
Qualquer tentativa de fraude conduzirá à anulação da prova para <u>todos</u> os intervenientes.
Respostas na folha de prova
Nas respostas múltiplas, as respostas erradas subtraem cotação.

1. Considere uma variável estocástica X∈{1,2,3} tal que P(X=1)=a/2, P(X=2)=1-a e P(x=3)=a/2, a∈[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) ≤ 2
      □ H(X) ≥ 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

2. Considere que Y = log2(2X+2). Assumindo que D() representa a distância KL, assinale as opções corretas:

   a) □ H(X,Y) ≥ 0 □ H(X,Y) ≤ 1
      □ H(Y) ≥ 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) ≤ 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) ≥ 0
□ D(X;Y)+D(Y;X) = 2H(Y) □ nenhuma das anteriores

3) Considere uma variável aleatória X que pode tomar os valores 0, 1, 2, ..., n, ... com probabilidade $1/e^2$, $2/e^2$, ..., $2^n/(n!e^2)$, ... Nestas condições observa-se que (Note que: $\sum_{n=1}^{\infty} ar^{-n} = \frac{a}{r-1}$, $\sum_{n=1}^{\infty} nr^{n} = \frac{r}{1-r^2}$, $\sum_{n=1}^{\infty} r^{n} = \frac{r}{1-r}$, $\sum_{i=0}^{\infty} \frac{x^n}{n!} = e^x$):

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

b) Um código de Huffman consegue representar os símbolos da sequência com:
□ menos de 5/3 bits/símb. □ nunca menos de 2/3 bits/símb.
□ Não é possível codificar com Huffman □ nenhuma das anteriores

4 – Assinale a(s) resposta(s) certa(s).

a) □ - A diferença entre um código Huffman e um código LZ77 é que o segundo é mais eficiente qualquer que seja o contexto.
□ - O código aritmético e o código Fano-Elias têm o mesmo princípio.
□ - A relação entre o LZW e o LZ77 é que o segundo não consegue capturar os padrões locais.
□ - Nenhuma das anteriores está correcta

b) □ - O código {0,10,110} é óptimo.
□ - O código {0,01,011} é instantâneo
□ - O código {0,01,011} é unicamente descodificável
□ - Nenhuma das anteriores está correcta

5 – Considere uma fonte de informação com alfabeto A={0,1,2}. Seja X a variável estocástica correspondente ao símbolo e Y a variável estocástica correspondente ao símbolo anterior numa cadeia de símbolos. Assuma que a distribuição conjunta P(X,Y) é a que se apresenta na tabela seguinte:

P(X,Y)	X=0	X=1	X=2
Y=0	1/9	1/9	1/9
Y=1	1/9	1/9	1/9
Y=2	1/9	1/9	1/9

a) É possível afirmar-se que a representação mais eficiente atingível por um código que permite codificar X é (arredondado às milésimas):
□ 1.102 bits/símb. □ 1.299 bits/símb.

□ 3.170 bits/símb. □ Nenhuma das anteriores

b) Aplicando um código de Huffman para codificar X, é possível garantir-se que o pior desempenho será:

□ 1.299 bits/símb. □ 2.299 bits/símb.
□ 4.170 bits/símb. □ Nenhuma das anteriores

c) Seja D() a distância KL. Observa-se que D(X,Y):

□ 1.499 □ 3.300
□ 1.399 □ Nenhuma das anteriores

d) Considere a sequência de X "001". A codificação da sequência usando um algoritmo aritmético poderá resultar na transmissão do seguinte código:

□ 0.1000 □ 0.1756
□ 0.5000 □ Nenhuma das anteriores

6 – Uma moeda de 2€ 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.

Observe que:

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

b) Qual a entropia quando a moeda é totalmente equilibrada?

7 - Considere uma fonte de informação pertencente ao dicionário S={1,2,3,4,5,6}. Assumindo que os símbolos são todos equiprováveis, indique a sequência de bits resultante da codificação da sequência "345" usando um código aritmético inteiro com 7 bits.

8 – Dada a fonte S e o alfabeto A, indique uma função do Python que permita a representação gráfica do respetivo histograma. Especifique os dois argumentos de entrada que a função requer para uma correta representação (considerando que para cada elemento do alfabeto é apresentada uma barra no histograma).
Como calculamos a entropia a partir dos valores (amplitudes) do histograma? Para responde a esta última questão apresente unicamente equações.

9 - Quer se converter uma imagem em cores numa imagem em escala de cinzas (Y). Para tal pretende-se utilizar o standard NTSC que aplica a seguinte transformação:

$$Y_{cinza}\left[i,j\right]=0.2978R\left[i,j\right]+0.5870G\left[i,j\right]+0.1140B\left[i,j\right]$$

onde R, G e B representam as cores vermelha, verde e azul, respetivamente. A seguir é apresentada parte de uma função implementada em Python que, dada a imagem em cores, retorna a imagem pretendida. Complete o código nos espaços indicados.

```
def rgb2gray(........................):

        import matplotlib.pyplot as plt
```

```
R, G, B = ..........................................................................................

Y = ..................................................................................................

plt.imshow(..........., cmap='gray')
plt.show()

return ...........
```