Universidade da Beira Interior

Curso: Engenharia Informática **Ano letivo: 2020-2021**

14333: Física e Informação

Teste 2 de Avaliação Contínua – possível resolução

1. Como $h(1) = 2\,\text{bits}$, temos $-\text{Log}_2\, p(1) = 2 \implies \text{Log}_2\, p(1) = -2 \implies p(1) = 2^{-2} = 0.25$. Então $p(6) = p(1) = 0.25$. Como $p(2) = p(5)$, $p(3) = p(4)$ e a soma das seis probabilidades é 1, temos $2(0.25) + 2p(2) + 2p(3) = 1 \implies p(2) + p(3) = 0.25$ [1].
Por outro lado, como $h(1) = h(6) = 2\,\text{bits}$, $h(2) = h(5)$, $h(3) = h(4)$ e a soma das seis surpresas é $16.91\,\text{bits}$, temos $2(2\,\text{bits}) + 2h(2) + 2h(3) = 16.91\,\text{bits} \implies h(2) + h(3) = 6.455\,\text{bits}$. Assim, $-\text{Log}_2\, p(2) - \text{Log}_2\, p(3) = 6.455 \implies \text{Log}_2[p(2)p(3)] = -6.455 \implies p(2)p(3) = 2^{-6.455} = 0.0114$. Portanto $p(3) = \frac{0.0114}{p(2)}$ [2]. Inserindo a expressão [2] na expressão [1] resulta
$$p(2) + \frac{0.0114}{p(2)} = 0.25 \implies [p(2)]^2 + 0.0114 = 0.25\,p(2) \implies [p(2)]^2 - 0.25\,p(2) + 0.0114 = 0\,.$$
Usando a fórmula resolvente, obtém-se $p(2) = \dfrac{0.25 \pm \sqrt{0.25^2 - 4(0.0114)}}{2} = \dfrac{0.25 \pm 0.13}{2} = \begin{cases} 0.06 \\ 0.19 \end{cases}$ e então $p(3) = 0.25 - p(2) = \begin{cases} 0.19 \\ 0.06 \end{cases}$. Conclui-se que as probabilidades dos números 2 e 3 são 0.19 e 0.06 (a ordem é irrelevante). Logo, a entropia associada ao lançamento deste dado será
$$H = -\frac{2(0.25 \ln 0.25 + 0.19 \ln 0.19 + 0.06 \ln 0.06)}{\ln 2}\,\text{bits} \approxeq 2.40\,\text{bits}.$$

2. Como a soma das quatro probabilidades é 1, temos $0.312 + \alpha + \beta + 0.224 = 1 \implies \alpha + \beta = 0.464$. Como $h(v_3) = 3h(v_2)$, temos $-\text{Log}_2\beta = 3(-\text{Log}_2\alpha) \implies \text{Log}_2\beta = \text{Log}_2\alpha^3 \implies \beta = \alpha^3$. Combinando as duas expressões obtidas, resulta $\alpha + \alpha^3 = 0.464$. Como a função $f(\alpha) = \alpha + \alpha^3$ é estritamente crescente, por inspeção é fácil descobrir que $\alpha = 0.4$ e $\beta = 0.4^3 = 0.064$. Assim, a distribuição de probabilidades da variável V é $\{0.312, 0.4, 0.064, 0.224\}$, pelo que a sua entropia (entropia inicial, antes de recebermos a informação) é
$$H_{\text{i}} = -\frac{0.312 \ln 0.312 + 0.4 \ln 0.4 + 0.064 \ln 0.064 + 0.224 \ln 0.224}{\ln 2}\,\text{bit} \approxeq 1.7903\,\text{bit}.$$
Quando nos dizem que o valor v_2 deixou de ser possível, a distribuição de probabilidades muda, e passamos a ter apenas três valores possíveis. Como a soma das três probabilidades correspondentes é $1 - 0.4 = 0.6$, as novas probabilidades serão
$$\mathcal{P}'(v_1) = \frac{0.312}{0.6} = 0.52\,, \qquad \mathcal{P}'(v_3) = \frac{0.064}{0.6} = \frac{8}{75} \qquad \text{e} \qquad \mathcal{P}'(v_4) = \frac{0.224}{0.6} = \frac{28}{75}\,.$$
Sendo a nova distribuição de probabilidades $\{0.52, 0, 8/75, 28/75\}$, a nova entropia (entropia final, após recebermos a informação) é
$$H_{\text{f}} = -\frac{0.52 \ln 0.52 + (8/75) \ln(8/75) + (28/75) \ln(28/75)}{\ln 2}\,\text{bit} \approxeq 1.3657\,\text{bit}.$$
(Note-se que o valor 0 pode ser ignorado, pois não contribui para a entropia.)
Logo, o valor I da informação é $I = H_{\text{i}} - H_{\text{f}} \approxeq 1.7903\,\text{bit} - 1.3657\,\text{bit} = 0.4246\,\text{bit} \approxeq 0.42\,\text{bit}$.

(*continua no verso desta folha*)

3. Usando a distribuição conjunta de probabilidades pode-se calcular a entropia conjunta:

$$H(M,N) = -\frac{0.23\ln 0.23 + 0.17\ln 0.17 + 0.21\ln 0.21 + 0.24\ln 0.24 + 0.15\ln 0.15}{\ln 2}\,\text{bits} \approxeq 2.29976\,\text{bits}.$$

(Ignora-se a probabilidade 0, pois ela não contribui para a entropia.)
As distribuições de probabilidade das variáveis M e N (distribuições marginais) podem ser obtidas somando os valores da tabela por linhas e por colunas, respetivamente:

$$\mathcal{P}(M=m_1) = 0.23 + 0.17 = 0.4\,;\ \mathcal{P}(M=m_2) = 0 + 0.21 = 0.21\,;\ \mathcal{P}(M=m_3) = 0.24 + 0.15 = 0.39\,.$$

$$\mathcal{P}(N=n_1) = 0.23 + 0 + 0.24 = 0.47\,;\ \mathcal{P}(N=n_2) = 0.17 + 0.21 + 0.15 = 0.53\,.$$

As respetivas entropias serão, portanto,

$$\begin{aligned} H(M) &= -\frac{0.4\ln 0.4 + 0.21\ln 0.21 + 0.39\ln 0.39}{\ln 2}\,\text{bit} \approxeq 1.53139\,\text{bit}. \\ H(N) &= -\frac{0.47\ln 0.47 + 0.53\ln 0.53}{\ln 2}\,\text{bit} \approxeq 0.99740\,\text{bit}. \end{aligned}$$

Então o valor da informação mútua partilhada pelas variáveis M e N é

$$I(M,N) = H(M) + H(N) - H(M,N) \approxeq (1.53139 + 0.99740 - 2.29976)\,\text{bit} = 0.22903\,\text{bit} \approxeq 0.229\,\text{bit}.$$

4. A matriz que representa o canal $I \to J$ é $[M_{JI}] = \begin{pmatrix} 0.4 & 0 & 0.2 & 0.3 \\ 0 & 1 & 0.8 & 0 \\ 0.6 & 0 & 0 & 0.7 \end{pmatrix}_{3\times 4}$ e a matriz que representa o canal $J \to K$ é $[M_{KJ}] = \begin{pmatrix} 0.9 & 1 & 0.5 \\ 0.1 & 0 & 0.5 \end{pmatrix}_{2\times 3}$, pelo que a matriz que representa o canal $I \to K$ será

$$[M_{KI}] = [M_{KJ}][M_{JI}] = \begin{pmatrix} 0.9 & 1 & 0.5 \\ 0.1 & 0 & 0.5 \end{pmatrix}_{2\times 3} \begin{pmatrix} 0.4 & 0 & 0.2 & 0.3 \\ 0 & 1 & 0.8 & 0 \\ 0.6 & 0 & 0 & 0.7 \end{pmatrix}_{3\times 4} = \begin{pmatrix} 0.66 & 1 & 0.98 & 0.62 \\ 0.34 & 0 & 0.02 & 0.38 \end{pmatrix}_{2\times 4}.$$

5. Como o canal trata os valores 1 e 3 da variável X exatamente da mesma maneira, há simetria nesses dois valores. Assim, na situação de informação mútua máxima, deverá ser $\mathcal{P}(X=1) = \mathcal{P}(X=3)$. Designando essas probabilidades por α, teremos $\alpha + \mathcal{P}(X=2) + \alpha = 1 \implies \mathcal{P}(X=2) = 1 - 2\alpha$. Podemos então usar $\{\alpha, 1-2\alpha, \alpha\}$ como distribuição de probabilidades de entrada. Logo

$$\mathcal{P}(Y=0) = \alpha(0.15) + (1-2\alpha)(0.5) + \alpha(0.85) = 0.15\alpha + 0.5 - \alpha + 0.85\alpha = 0.5\,,$$

$$\mathcal{P}(Y=1) = \alpha(0.85) + (1-2\alpha)(0.5) + \alpha(0.15) = 0.85\alpha + 0.5 - \alpha + 0.15\alpha = 0.5\,.$$

A distribuição de probabilidades de saída será, portanto, $\{0.5, 0.5\}$, ou seja, uma distribuição uniforme. Para calcular a informação mútua usamos a expressão $I(X,Y) = H(Y) - H(Y|X)$. Como a distribuição de Y é uniforme, $H(Y) = (\text{Log}_2 2)\,\text{bit} = 1\,\text{bit}$. Para calcular $H(Y|X)$, há que considerar três cenários:
Cenário $X=1$: $H(Y|X=1) = H(0.15, 0.85) = -\frac{0.15\ln 0.15 + 0.85\ln 0.85}{\ln 2} \approxeq 0.61\,\text{bit}$;
Cenário $X=2$: $H(Y|X=2) = H(0.5, 0.5) = (\text{Log}_2 2)\,\text{bit} = 1\,\text{bit}$;
Cenário $X=3$: $H(Y|X=3) = H(0.85, 0.15) = -\frac{0.85\ln 0.85 + 0.15\ln 0.15}{\ln 2} \approxeq 0.61\,\text{bit}$.
Então

$$H(Y|X) \approxeq \alpha(0.61\,\text{bit}) + (1-2\alpha)(1\,\text{bit}) + \alpha(0.61\,\text{bit}) = (1.22\alpha + 1 - 2\alpha)\,\text{bit} = (1 - 0.78\alpha)\,\text{bit}.$$

Portanto $I(X,Y) \approxeq [1 - (1 - 0.78\alpha)]\,\text{bit} = 0.78\alpha\,\text{bit}$. A informação mútua I será máxima para o maior valor possível de α. Devido à condição $1 - 2\alpha \geq 0$, vemos que $\alpha \leq 0.5$, pelo que o valor máximo de α é 0.5. Assim, a capacidade do canal será

$$C(X,Y) = [I(X,Y)]_{\text{máx}} \approxeq 0.78(0.5)\,\text{bit} = 0.39\,\text{bit}.$$