Informação Mútua e Propriedades da Função Entropia

Teoria da Informação - AULA 04

Profª. Verusca Severo

Universidade de Pernambuco
Escola Politécnica de Pernambuco

25 de junho de 2021

Relembrando Entropia Conjunta e Condicional

- Quando estudamos sistemas que são descritos por duas ou mais variáveis uma quantidade importante a ser definida é a **entropia conjunta** e **entropia condicional**.
- Seja $X = \{x_1, x_2, \ldots, x_k\}$ e $Y = \{y_1, y_2, \ldots, y_l\}$ um par de variáveis aleatórias.
  - A **entropia conjunta** é:

$$H(X, Y) = \sum_{i=1}^{k} \sum_{j=1}^{l} p(X = x_i, Y = y_j) \log_2 \frac{1}{p(X = x_i, Y = y_j)}$$

  - A **entropia condicional** de X dado que o evento $Y = y_q$ ocorreu é:

$$H(X/Y = y_q) = \sum_{i=1}^{k} P(X = x_i/Y = y_q) \log_2 \left(\frac{1}{P(X = x_i/Y = y_q)} \right)$$

Relembrando Entropia Conjunta e Condicional

- Seja $X = \{x_1, x_2, \ldots, x_k\}$ e $Y = \{y_1, y_2, \ldots, y_l\}$ um par de variáveis aleatórias.
  - A **entropia condicional média** de X dada a variável Y é:

$$H(X/Y) = \sum_{j=1}^{m} P(Y = y_j) H(X/Y = y_j)$$

- **Regra da cadeia (Teorema):** A entropia verifica a seguinte propriedade

$$H(X, Y) = H(X) + H(Y/X) = H(Y) + H(X/Y).$$

Entropia Condicional

- **Condicionamento reduz a entropia (Teorema):** Para quaisquer duas variáveis aleatórias X e Y a expressão

$$H(X/Y) \leq H(X)$$

é verificada a igualdade se e somente se X e Y forem estatisticamente independentes.

- O teorema acima surge "intuitivamente" do fato de que o conhecimento de Y reduz, em geral, a incerteza que se tem sobre X.

Informação Mútua

- Quando duas variáveis X e Y estão correlacionadas, o conhecimento de uma variável reduz a incerteza associada à outra variável.

- Para a pergunta: "Quanta informação a variável aleatória Y fornece sobre a variável aleatória X?"
  - a resposta de Shannon foi: **"é o valor pelo qual Y reduz a incerteza sobre X".**

  - A redução da incerteza, que está diretamente relacionada com a redução da quantidade de informação devido a correlação, chamamos de **informação mútua**.

- **Definição:** A informação mútua entre as variáveis aleatórias discretas X e Y é por definição a quantidade:

$$I(X,Y) = H(X) - H(X/Y)$$

ou

$$I(Y,X) = H(Y) - H(Y/X)$$

Informação Mútua

- A denominação **informação mútua** é bastante apropriada pelo fato de que:
$$I(X, Y) = I(Y, X).$$
- Para tal, inicialmente $H(X, Y)$ é expandido de duas maneiras
$$H(X, Y) = H(X) + H(Y/X) \tag{1}$$
e
$$H(X, Y) = H(Y) + H(X/Y) \tag{2}$$
- Fazendo (1) - (2), temos:
$$0 = H(X) + H(Y/X) - H(Y) - H(X/Y) \tag{3}$$

Informação Mútua

- Organizando (3), temos:

$$H(Y) - H(Y/X) = H(X) - H(X/Y)$$
$$I(Y, X) = I(X, Y)$$

- A igualdade acima significa que a quantidade de informação provida pela variável aleatória X sobre a variável aleatória Y é exatamente a mesma quantidade de informação provida pela variável aleatória Y sobre a variável aleatória X.

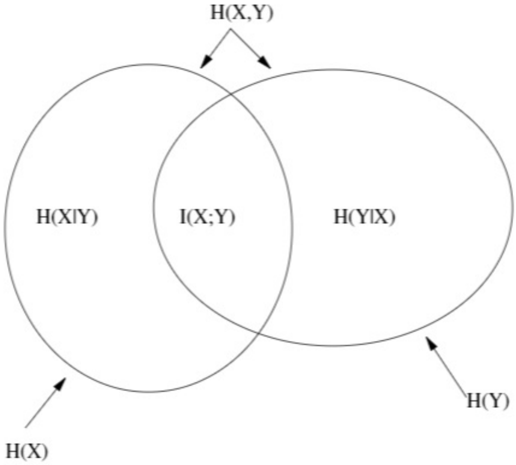
- Trata-se portanto de uma relação simétrica entre as variáveis X e Y.

Relações entre Informação Mútua e Entropia:

- $I(X, Y) = I(Y, X)$
- $I(X, Y) = H(X) - H(X/Y)$
- $H(X) = H(X/Y) + I(X, Y)$
- $I(Y, X) = H(Y) - H(Y/X)$
- $H(Y) = H(Y/X) + I(Y, X)$
- $H(X, Y) = H(X) + H(Y) - I(X, Y)$

Relação entre Informação Mútua e Entropia:

- A Figura ilustra essas relações usando um diagrama de Venn



Exercício

1. Suponha que o vetor aleatório $[X_1, X_2, X_3]$ assume os valores $[0, 0, 0].[0, 1, 1], [1, 0, 1]$ e $[1, 1, 0]$, cada um deles com probabilidade $\frac{1}{4}$. Calcule:
   a) $H(X_1)$, $H(X_2)$ e $H(X_3)$
   b) $H(X_1, X_2, X_3)$
   c) $H(X_2/X_1)$
   d) $H(X_3/X_1X_2)$
   e) Verifique se a expressão
   $H(X_1, X_2, X_3) = H(X_1) + H(X_2/X_1) + H(X_3/X_1X_2)$ é satisfeita.

Exercício

- **SOLUÇÃO: a)** $H(X_1)$, $H(X_2)$ e $H(X_3)$
  - Determinando $H(X_1)$:

$$P(X_1=0)=P(X_1=0,X_2=0,X_3=0)+P(X_1=0,X_2=1,X_3=1)$$

$$P(X_1=0)=1/4+1/4=1/2 \quad \Rightarrow \quad P(X_1=1)=1/2$$

Logo:

$$H(X_1)=\frac{1}{2}\log_2\left[\frac{1}{1/2}\right]+\frac{1}{2}\log_2\left[\frac{1}{1/2}\right]=1\ bit$$

  - Determinando $H(X_2)$:

$$P(X_2=0)=P(X_1=0,X_2=0,X_3=0)+P(X_1=1,X_2=0,X_3=1)$$

$$P(X_2=0)=1/4+1/4=1/2 \quad \Rightarrow \quad P(X_2=1)=1/2$$

Logo:

$$H(X_2)=\frac{1}{2}\log_2\left[\frac{1}{1/2}\right]+\frac{1}{2}\log_2\left[\frac{1}{1/2}\right]=1\ bit$$

- **SOLUÇÃO: a)** $H(X_1)$, $H(X_2)$ e $H(X_3)$
  - Determinando $H(X_3)$:

$$P(X_3 = 0) = P(X_1 = 0, X_2 = 0, X_3 = 0) + P(X_1 = 1, X_2 = 1, X_3 = 0)$$

$$P(X_3 = 0) = 1/4 + 1/4 = 1/2 \quad \Rightarrow \quad P(X_3 = 1) = 1/2$$

Logo:

$$H(X_3) = \frac{1}{2}\log_2\left[\frac{1}{1/2}\right] + \frac{1}{2}\log_2\left[\frac{1}{1/2}\right] = 1 \ bit$$

Exercício

- **SOLUÇÃO: b)** $H(X_1, X_2, X_3)$
  - Usando a distribuição de prob. conjunta, temos que

$$H(X_1, X_2, X_3) = 4 \times \left(\frac{1}{4} \log_2 \left[\frac{1}{1/4} \right] \right)$$

$$H(X_1, X_2, X_3) = 4 \times \left(\frac{1}{4} \log_2 \left[\frac{1}{1/4} \right] \right)$$

$$H(X_1, X_2, X_3) = \log_2 \left[\frac{1}{1/4} \right]$$

$$H(X_1, X_2, X_3) = \log_2 [4]$$

$$H(X_1, X_2, X_3) = 2 \ bits$$

- **SOLUÇÃO: c)** $H(X_2|X_1) = ?$
$H(X_2|X_1) = P(X_1 = 0) \times H(X_2|X_1 = 0) + P(X_1 = 1) \times H(X_2|X_1 = 1)$
  - **Determinando $H(X_2|X_1 = 0)$:**
  Fazendo $P(X_2 = 0|X_1 = 0) = p \quad \Rightarrow \quad P(X_2 = 1|X_1 = 0) = 1 - p$:

$$H(X_2|X_1 = 0) = p \log_2 \left[\frac{1}{p}\right] + (1-p) \log_2 \left[\frac{1}{1-p}\right]$$

  Precisamos determinar $P(X_2 = 0|X_1 = 0)$:

$$P(X_2 = 0|X_1 = 0) = \frac{P(X_1 = 0, X_2 = 0, X_3 = 0)}{P(X_1 = 0)} = \frac{1/4}{1/2} = 1/2 = p$$

$$P(X_2 = 1|X_1 = 0) = 1 - P(X_2 = 0|X_1 = 0) = 1/2 = (1-p)$$

  Logo:

$$H(X_2|X_1 = 0) = \frac{1}{2} \log_2 \left[\frac{1}{1/2}\right] + \frac{1}{2} \log_2 \left[\frac{1}{1/2}\right] = 1 \quad bit$$

Exercício

- **SOLUÇÃO: c)** $H(X_2|X_1) = ?$
  $H(X_2|X_1) = P(X_1 = 0) \times H(X_2|X_1 = 0) + P(X_1 = 1) \times H(X_2|X_1 = 1)$
  - **Determinando $H(X_2|X_1 = 1)$:** Fazendo
    $P(X_2 = 0|X_1 = 1) = p \quad \Rightarrow \quad P(X_2 = 1|X_1 = 1) = 1 - p$:

    $$H(X_2|X_1 = 1) = p\log_2\left[\frac{1}{p}\right] + (1-p)\log_2\left[\frac{1}{1-p}\right]$$

    Precisamos determinar $P(X_2 = 0|X_1 = 1)$:

    $$P(X_2 = 0|X_1 = 1) = \frac{P(X_1 = 1, X_2 = 0, X_3 = 1)}{P(X_1 = 1)} = \frac{1/4}{1/2} = 1/2 = p$$

    $$P(X_2 = 1|X_1 = 1) = 1 - P(X_2 = 0|X_1 = 1) = 1/2 = (1-p)$$

    Logo:

    $$H(X_2|X_1 = 1) = \frac{1}{2}\log_2\left[\frac{1}{1/2}\right] + \frac{1}{2}\log_2\left[\frac{1}{1/2}\right] = 1 \quad bit$$

Exercício

- **SOLUÇÃO: c)** $H(X_2|X_1) = ?$
  - Logo:

$$H(X_2|X_1) = P(X_1 = 0) \times H(X_2|X_1 = 0) + P(X_1 = 1) \times H(X_2|X_1 = 1)$$

$$H(X_2|X_1) = \frac{1}{2} \times 1 + \frac{1}{2} \times 1$$

$$H(X_2|X_1) = 1 \quad bit$$

Exercício

- **SOLUÇÃO: d)** $H(X_3|X_1X_2) =?$

  - Temos que:

$$\begin{aligned} H(X_3|X_1X_2) = & P(X_1X_2 = 00)H(X_3|X_1X_2 = 00) + \\ & P(X_1X_2 = 01)H(X_3|X_1X_2 = 01) + \\ & P(X_1X_2 = 10)H(X_3|X_1X_2 = 10) + \\ & P(X_1X_2 = 11)H(X_3|X_1X_2 = 11) \end{aligned}$$

  - Determinando $H(X_3|X_1X_2 = 00)$:

$$\begin{aligned} H(X_3|X_1X_2 = 00) = & P(X_3 = 0|X_1X_2 = 00)\log_2\left[\frac{1}{P(X_3 = 0|X_1X_2 = 00)}\right] \\ & + P(X_3 = 1|X_1X_2 = 00)\log_2\left[\frac{1}{P(X_3 = 1|X_1X_2 = 00)}\right] \end{aligned}$$

Exercício

- **SOLUÇÃO: d)** $H(X_3|X_1X_2) = ?$

1. Determinando $H(X_3|X_1X_2 = 00)$:

$$H(X_3|X_1X_2 = 00) = P(X_3 = 0|X_1X_2 = 00)\log_2\left[\frac{1}{P(X_3 = 0|X_1X_2 = 00)}\right]$$
$$+P(X_3 = 1|X_1X_2 = 00)\log_2\left[\frac{1}{P(X_3 = 1|X_1X_2 = 00)}\right]$$

- Determinando $P(X_3 = 0|X_1X_2 = 00)$:

$$P(X_3 = 0|X_1X_2 = 00) = \frac{P(X_1X_2X_3 = 000)}{P(X_1X_2 = 00)} = \frac{1/4}{1/4} = 1$$

$$P(X_3 = 1|X_1X_2 = 00) = 1 - P(X_3 = 0|X_1X_2 = 00) = 0$$

- Logo:

$$H(X_3|X_1X_2 = 00) = 0 \ bit$$

Exercício

- **SOLUÇÃO: d)** $H(X_3|X_1X_2) = ?$

  2. Determinando $H(X_3|X_1X_2 = 01)$:

$$H(X_3|X_1X_2 = 01) = P(X_3 = 0|X_1X_2 = 01)\log_2\left[\frac{1}{P(X_3 = 0|X_1X_2 = 01)}\right]$$
$$+P(X_3 = 1|X_1X_2 = 01)\log_2\left[\frac{1}{P(X_3 = 1|X_1X_2 = 01)}\right]$$

  - Determinando $P(X_3 = 0|X_1X_2 = 01)$:

$$P(X_3 = 0|X_1X_2 = 01) = \frac{P(X_1X_2X_3 = 010)}{P(X_1X_2 = 01)} = \frac{0}{1/4} = 0$$

$$P(X_3 = 1|X_1X_2 = 01) = 1 - P(X_3 = 0|X_1X_2 = 01) = 1$$

  - Logo:

$$H(X_3|X_1X_2 = 01) = 0 \ bit$$

- **SOLUÇÃO: d)** $H(X_3|X_1X_2) = ?$

3. Determinando $H(X_3|X_1X_2 = 10)$:

$$H(X_3|X_1X_2 = 10) = P(X_3 = 0|X_1X_2 = 10)\log_2\left[\frac{1}{P(X_3 = 0|X_1X_2 = 10)}\right] + P(X_3 = 1|X_1X_2 = 10)\log_2\left[\frac{1}{P(X_3 = 1|X_1X_2 = 10)}\right]$$

- Determinando $P(X_3 = 0|X_1X_2 = 01)$:

$$P(X_3 = 0|X_1X_2 = 10) = \frac{P(X_1X_2X_3 = 100)}{P(X_1X_2 = 10)} = \frac{0}{1/4} = 0$$

$$P(X_3 = 1|X_1X_2 = 10) = 1 - P(X_3 = 0|X_1X_2 = 10) = 1$$

- Logo:

$$H(X_3|X_1X_2 = 10) = 0 \ bit$$

Exercício

- **SOLUÇÃO: d)** $H(X_3|X_1X_2) = ?$

4. Determinando $H(X_3|X_1X_2 = 11)$:

$$H(X_3|X_1X_2 = 11) = P(X_3 = 0|X_1X_2 = 11)\log_2\left[\frac{1}{P(X_3 = 0|X_1X_2 = 11)}\right]$$
$$+P(X_3 = 1|X_1X_2 = 11)\log_2\left[\frac{1}{P(X_3 = 1|X_1X_2 = 11)}\right]$$

   - Determinando $P(X_3 = 0|X_1X_2 = 11)$:

$$P(X_3 = 0|X_1X_2 = 11) = \frac{P(X_1X_2X_3 = 110)}{P(X_1X_2 = 11)} = \frac{1/4}{1/4} = 1$$

$$P(X_3 = 1|X_1X_2 = 11) = 1 - P(X_3 = 0|X_1X_2 = 11) = 0$$

   - Logo:

$$H(X_3|X_1X_2 = 11) = 0 \ bit$$

Exercício

- **SOLUÇÃO: d)** $H(X_3|X_1X_2) = ?$

  - Substituindo as entropias obtidas em:

$$\begin{aligned} H(X_3|X_1X_2) = \; & P(X_1X_2 = 00)H(X_3|X_1X_2 = 00) + \\ & P(X_1X_2 = 01)H(X_3|X_1X_2 = 01) + \\ & P(X_1X_2 = 10)H(X_3|X_1X_2 = 10) + \\ & P(X_1X_2 = 11)H(X_3|X_1X_2 = 11) \end{aligned}$$

  - Temos que:

$$H(X_3|X_1X_2) = 0 \;\; bit$$

Exercício

- **SOLUÇÃO: e)** Verifique se a expressão $H(X_1, X_2, X_3) = H(X_1) + H(X_2/X_1) + H(X_3/X_1X_2)$ é satisfeita.
  - Substituindo as entropias obtidas em:

$$H(X_1, X_2, X_3) = H(X_1) + H(X_2/X_1) + H(X_3/X_1X_2)$$

  - Temos que:

$$2 = 1 + 1 + 0$$

$$2 = 2$$

  - Logo, a expressão é satisfeita!