# Final 2 2021:

## Problema 1:

**Enunciado:** Si es posible, encuentre ejemplos de variables $X$ e $Y$ tales que

(a) $H(X) = 1$ bit, $H(Y) = 1$ bit, $I(X;Y) = 0$.

(b) $H(X) = 2$ bit, $H(Y) = 1$ bit, $I(X;Y) = 1$ bit.

(c) $H(X) = 2$ bit, $H(Y) = 2$ bit, $I(X;Y) = 1$ bit.

(d) $H(X) = 1$ bit, $H(Y) = 2$ bit, $I(X;Y) = 2$ bit.

(e) $H(X) = 2$ bit, $H(Y) = 3$ bit, $I(X;Y) = 1$ bit.

Encontrar un ejemplo significa dar los conjuntos $\mathcal{A}_x$y$\mathcal{A}_Y$ donde $X$ e $Y$ toman valores, y también dar explícitamente las distribuciones de probabilidad $p(x), p(y)$ y $p(x,y), \forall x \in \mathcal{A}_X$y$\forall y \in \mathcal{A}_Y$. Si no existe ningún par de variables aleatorias $X$ e $Y$ que cumpla con lo requerido por el enunciado, demuestre tal no existencia.

**Resolucion:**

### Preámbulo

Podemos verificar que las probabilidades condicionales están bien definidas usando:

$$\sum_x p(x \mid y) = 1$$

### Inciso a

Como $I(X;Y) = 0 \Leftrightarrow XeY$ son indep. $\Rightarrow P(x,y) = P(x)P(y)$

Se puede tener $H(X) = 1$ `bit` si $p(x) = \frac{1}{2}, \forall x \in A_X$, con $|A_X| = 2$.

Lo mismo es cierto para $Y$.

### Inciso b

Dados:

$H(X) = 2$ `bit`, $\quad H(Y) = 1$ `bit`, $\quad I(X;Y) = 1$ `bit`

Por simplicidad se propone:

$$|A_Y| = 2, |A_X| = 4$$

$$x = \Big\{0, 1, \text{con probabilidad } \frac{1}{2}$$

$$y = \Big\{0, 1, 2, 3, \text{con probabilidad } \frac{1}{4}$$

Es sabido que:

$$0 \leq I(X;Y) \leq \min(H(X), H(Y))$$

$I$ es máxima lo cual sugiere que se ahorra el máximo # de preguntas posible. Esto descarta que ambas variables sean independientes. Si uno propone $X$ e $Y$ relacionados por $f$:

En forma matricial:

| $p(x \mid y)$ | $y=0$ | $y=1$ |
|---|---|---|
| $x=0$ | $1/2$ | $0$ |
| $x=1$ | $1/2$ | $0$ |
| $x=2$ | $0$ | $1/2$ |
| $x=3$ | $0$ | $1/2$ |

La conjunta: $p(x,y) = p(x \mid y)p(y)$

| $p(x,y)$ | $y=0$ | $y=1$ |
|---|---|---|
| $x=0$ | $1/8$ | $0$ |
| $x=1$ | $1/8$ | $0$ |
| $x=2$ | $0$ | $1/8$ |
| $x=3$ | $0$ | $1/8$ |

Calculando $H(X \mid Y)$:

$$\begin{aligned} H(X \mid Y) &= -\sum_y p(y) \sum_x p(x \mid y) \log_2 p(x \mid y) \\ &= -\sum_y p(y) \sum_x a_{xy} = -\sum_y \frac{1}{2} \sum_x a_{xy} \\ &= -\frac{1}{2} \sum_y \sum_x a_{xy} = 1\texttt{bit} \end{aligned}$$

Usando que $a_{xy}$

| $a_{xy}$ | $y=0$ | $y=1$ |
|---|---|---|
| $x=0$ | $-1/2$ | $0$ |
| $x=1$ | $-1/2$ | $0$ |
| $x=2$ | $0$ | $-1/2$ |
| $x=3$ | $0$ | $-1/2$ |

$I(X;Y) = H(X) - H(X \mid Y) = 2\texttt{bit} - 1\texttt{bit} = 1\texttt{bit}$

Las variables $X, Y$ definidas anteriormente cumplen todas las condiciones que se piden.

**Inciso c**

Dados:

$H(X) = 2\ \texttt{bit}, \quad H(Y) = 2\ \texttt{bit}, \quad I(X;Y) = 1\ \texttt{bit}$

Que pueden obtenerse con variables uniformemente distribuidas:

$$|A_Y| = 4, |A_X| = 4$$

Es sabido que:

$$0 \leq I(X;Y) \leq \min(H(X), H(Y))$$

Lo cual no contradice que $I(X;Y) = 1\ \texttt{bit}$.

Se ahorra un numero de preguntas. Por lo que las variables no son totalmente independientes, suponiendo una dependencia tipo:

$$y = \Big\{0, 1, 2, 3, \text{con probabilidad } \frac{1}{4}$$

$$x = \Big\{0, 1, 2, 3, \text{con probabilidad } \frac{1}{4}$$

| $p(x \mid y)$ | $y=0$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|---|
| $x=0$ | $1/2$ | $1/2$ | $0$ | $0$ |
| $x=1$ | $1/2$ | $1/2$ | $0$ | $0$ |
| $x=2$ | $0$ | $0$ | $1/2$ | $1/2$ |
| $x=3$ | $0$ | $0$ | $1/2$ | $1/2$ |

| $p(x,y)$ | $y=0$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|---|
| $x=0$ | $1/8$ | $1/8$ | $0$ | $0$ |
| $x=1$ | $1/8$ | $1/8$ | $0$ | $0$ |
| $x=8$ | $0$ | $0$ | $1/8$ | $1/8$ |
| $x=3$ | $0$ | $0$ | $1/8$ | $1/8$ |

$$\begin{aligned} H(X \mid Y) &= -\sum_y p(y) \sum_x p(x \mid y) \log_2 p(x \mid y) \\ &= -\sum_y p(y) \sum_x a_{xy} = -\sum_y p(y) \sum_x a_{xy} \\ &= -\sum_y \frac{1}{4} \sum_x a_{xy} = -\frac{1}{4} \sum_y \sum_x a_{xy} \\ &= 1\texttt{bit} \end{aligned}$$

Usando que:

| $a_{xy}$ | $y=0$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|---|
| $x=0$ | $-1/2$ | $-1/2$ | $0$ | $0$ |
| $x=1$ | $-1/2$ | $-1/2$ | $0$ | $0$ |
| $x=2$ | $0$ | $0$ | $-1/2$ | $-1/2$ |
| $x=3$ | $0$ | $0$ | $-1/2$ | $-1/2$ |

$I(X;Y) = H(X) - H(X \mid Y) = 2\texttt{bit} - 1\texttt{bit} = 1\texttt{bit}$

**Inciso d**

Dados:

$H(X) = 1\ \texttt{bit}, \quad H(Y) = 2\ \texttt{bit}, \quad I(X;Y) = 2\ \texttt{bit}$

Que pueden obtenerse con variables uniformemente distribuidas:

$$|A_Y| = 2, |A_X| = 4$$

Es sabido que:

$$0 \leq I(X;Y) \leq \min(H(X), H(Y))$$

Lo cual contradice que $I(X;Y) = 2\ \texttt{bit}$, por lo que no es posible tener variables $X$, $Y$ que cumplan estos requisitos.

**Inciso e**

Dados:

$H(X) = 2\ \texttt{bit}, \quad H(Y) = 3\ \texttt{bit}, \quad I(X;Y) = 1\ \texttt{bit}$

Que pueden obtenerse con variables uniformemente distribuidas:

$$|A_Y| = 2, |A_X| = 8$$

Es sabido que:

$$0 \leq I(X;Y) \leq \min(H(X), H(Y))$$

Lo cual no contradice que $I(X;Y) = 1\ \texttt{bit}$.

$$y = \Big\{0, 1, 2, 3, 4, 5, 6, 7,\ \text{con probabilidad } \frac{1}{8}$$

$$x = \left\{ 0, 1, 2, 3, \text{con probabilidad } \frac{1}{4} \right.$$

| $p(x \mid y)$ | $y=0$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|---|
| $x=0$ | $1/2$ | $0$ | $0$ | $0$ |
| $x=1$ | $1/2$ | $0$ | $0$ | $0$ |
| $x=2$ | $0$ | $1/2$ | $0$ | $0$ |
| $x=3$ | $0$ | $1/2$ | $0$ | $0$ |
| $x=4$ | $0$ | $0$ | $1/2$ | $0$ |
| $x=5$ | $0$ | $0$ | $1/2$ | $0$ |
| $x=6$ | $0$ | $0$ | $0$ | $1/2$ |
| $x=7$ | $0$ | $0$ | $0$ | $1/2$ |

La probabilidad conjunta:

| $p(x, y)$ | $y=0$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|---|
| $x=0$ | $1/16$ | $0$ | $0$ | $0$ |
| $x=1$ | $1/16$ | $0$ | $0$ | $0$ |
| $x=2$ | $0$ | $1/16$ | $0$ | $0$ |
| $x=3$ | $0$ | $1/16$ | $0$ | $0$ |
| $x=4$ | $0$ | $0$ | $1/16$ | $0$ |
| $x=5$ | $0$ | $0$ | $1/16$ | $0$ |
| $x=6$ | $0$ | $0$ | $0$ | $1/16$ |
| $x=7$ | $0$ | $0$ | $0$ | $1/16$ |

Calculando $H(X \mid Y)$:

$$\begin{aligned} H(X \mid Y) &= -\sum_y p(y) \sum_x p(x \mid y) \log_2 p(x \mid y) \\ &= -\sum_y p(y) \sum_x a_{xy} = -\sum_y p(y) \sum_x a_{xy} \\ &= -\sum_y \frac{1}{4} \sum_x a_{xy} = -\frac{1}{4} \sum_y \sum_x a_{xy} \\ &= 1\texttt{bit} \end{aligned}$$

usando que:

| $a_{xy}$ | $y=0$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|---|
| $x=0$ | $-1/2$ | $0$ | $0$ | $0$ |
| $x=1$ | $-1/2$ | $0$ | $0$ | $0$ |
| $x=2$ | $0$ | $-1/2$ | $0$ | $0$ |
| $x=3$ | $0$ | $-1/2$ | $0$ | $0$ |
| $x=4$ | $0$ | $0$ | $-1/2$ | $0$ |
| $x=5$ | $0$ | $0$ | $-1/2$ | $0$ |
| $x=6$ | $0$ | $0$ | $0$ | $-1/2$ |
| $x=7$ | $0$ | $0$ | $0$ | $-1/2$ |

Por lo que se puede garantizar que:

$$I(X; Y) = H(X) - H(X \mid Y) = 2\texttt{bit} - 1\texttt{bit} = 1\texttt{bit}$$

## Problema 2:

**Enunciado:** Una fuente genera símbolos $X \in \mathcal{A}_X = \{0, 1, 2\}$, con probabilidades $p_0 = 1/2, p_1 = 1/4, p_2 = 1/4,$ yse los codifica con un código aritmético binario con símbolos $Y \in \mathcal{A}_Y = \{0, 1\}$. Decodifique la secuencia que comienza con 11001000111011010111001011100111111... hasta donde le sea posible. Justifique.

**Resolucion:**

11001000111011010111001011100111111

Propongo hacer huffman:

```
1/2 1/4 1/4
 |   |   |
 |   \0 1/
 |     |
1/2   1/2
 |     |
\0    1/
  \   /
   \ /
    1
```

Con esta propuesta:

$$\begin{aligned} X = 0&, \; C(0) = 0 \\ X = 1&, \; C(1) = 10 \\ X = 2&, \; C(2) = 11 \end{aligned}$$

11 0 0 10 0 0 11 10 11 0 10 11 10 0 10 11 10 0 11 11 11

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | 0 | 0 | 10 | 0 | 0 | 11 | 10 | 11 | 0 | 10 | 11 | 10 | ... |
| 2 | 0 | 0 | 1 | 0 | 0 | 2 | 1 | 2 | 0 | 1 | 2 | 1 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 10 | 11 | 10 | 0 | 11 | 11 | 11 |
| 0 | 1 | 2 | 1 | 0 | 2 | 2 | 2 |

## Problema 3:

**Enunciado:** Considere las 10 urnas $(a, b, \ldots, j)$ de la figura. Se elige una urna al azar. La variable $X$ es la urna elegida. De esa urna, se extrae una pelota. La variable $Y$ es el color de la pelota extraída. Considere el canal $X \to Y$ que mapea la urna elegida con la pelota extraída. Calcule la capacidad del canal.

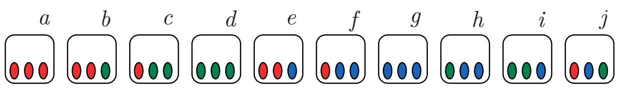


Figure 1: Urnas y pelotas

**Resolucion:**

Propongo:

$$p(x) = \begin{cases} a \text{ prob. } \dfrac{1}{3} \\ d \text{ prob. } \dfrac{1}{3} \\ g \text{ prob. } \dfrac{1}{3} \\ c.o.c \text{ prob. } 0 \end{cases}$$

esto me elimina filas en $P(X \mid Y)$ resultando

| $p(x \mid y)$ | $y = r$ | $y = g$ | $y = b$ |
|---|---|---|---|
| $x = a$ | 1 | 0 | 0 |
| $x = d$ | 0 | 1 | 0 |
| $x = g$ | 0 | 0 | 1 |

Uso otra notacion para los $x$ e $y$, que son $A_x = \{x_1, x_2, x_3\}, A_y = \{y_1, y_2, y_3\}$

para que la probabilidad condicional quede:

| $p(x \mid y)$ | $y = 1$ | $y = 2$ | $y = 3$ |
|---|---|---|---|
| $x = 1$ | 1 | 0 | 0 |
| $x = 2$ | 0 | 1 | 0 |
| $x = 3$ | 0 | 0 | 1 |

$= \delta_{xy}$

Calculo I(X;Y):

$$I(X;Y) = \sum_{xy} p(x)p(x \mid y) \log \frac{p(x \mid y)}{\sum_{x'} p(x')p(y \mid x')}$$

es uniforme en el alfabeto reducido $A'_x$, los otros valores no ocurren

$$\begin{aligned} I(X;Y) &= \sum_{x,y} \frac{1}{3}\delta_{x,y} \log \delta_{x,y} - \sum_{x,y} \frac{1}{3}\delta_{x,y} \log \frac{1}{3}\delta_{y,y} \\ &= -\sum_{x,y} \frac{1}{3}\delta_{x,y} \log \frac{1}{3} = -\frac{1}{3} \log \frac{1}{3} \sum_{x,y} \delta_{x,y} \\ &= -\log \frac{1}{3} = \log_2 3 \end{aligned}$$

Esta Informacion mutua es igual a $log|A_X|$, por esto:

$$C = log_2 3$$

## Problema 4:

**Enunciado:** En un canal binario simétrico $X \to Y$ con probabilidad de error de bit $p(y = 1 \mid x = 0) = p(y = 0 \mid x = 1) = q$ se transmite un código conformado por palabras clave de longitud $n$. Sea $Z$ una variable aleatoria binaria que representa si la cadena de $n$ digitos fue transmitida con o sin error. Se define un estimador $\hat{q}(z)$ que suponemos no sesgado, que estima la probabilidad de error de bit $q$ con que opera el canal a partir de una medición de $z$. Encuentre el minimo error cuadrático medio que puede tener el estimador $\hat{q}(z)$." Como se modifica la respuesta si consideramos un estimador no sesgado $\hat{q}(z_1, \ldots, z_k)$ que opera sobre $k$ muestras independientes de la variable binaria $Z$?

**Resolucion:**

EL error cuadratico minimo (por la cota de cramer Rao) es :

$$E^2(q) = \frac{1}{J(q)}$$

$$error = \begin{cases} \text{no hay error } ,1 - q \\ \text{hay error } ,q \end{cases}$$

Para un codigo de longitud $n$ la variable $Z$, con una probabilidad de tener error $q$, tenemos una distribucion binomial:

$$z \in \{1, ..., n\}, P(z \mid q) = \frac{n!}{z!(n-z)!} q^z (1-q)^{n-z}$$

con una información de fisher:

$$J(q) = \frac{n}{q(1-q)}$$

Con lo que:

$$E^2_{\min}(q) = \frac{q(1-q)}{n}$$

Como la informaicon de fisher es aditiva y tengo k muestras:

$$J_k(q) = kJ_1(q) = \frac{nk}{q(1-q)}$$

Con lo que:

$$E^2_{\min}(q) = \frac{q(1-q)}{nk}$$

Las propiedades de la informacion de fisher me permiten tener una idea de cual es el error cuadratico medio de un “buen” estimador sin tener que realizar propuestas de estimadores y compararlas.