Comunicaciones Digitales

Práctico 7

Teoría de la Información: información mutua y canal binario simétrico

Cada ejercicio comienza con un símbolo el cuál indica su dificultad de acuerdo a la siguiente escala: ✦ básica, ★ media, ✷ avanzada, y ✹ difícil.

✦ Ejercicio 1

Demostrar que si X e Y son símbolos discretos, aleatorios e independientes, se cumple que su entropía conjunta es la suma de sus entropías.

✦ Ejercicio 2

Sean X e Y símbolos discretos, aleatorios e independientes. Las entropías condicional $H\{X|Y\}$ y conjunta $H\{X;Y\}$ se definen como:

$$\begin{aligned} H\{X|Y\} &= E\{\log_2 \tfrac{1}{p(x|y)}\} \\ H\{X;Y\} &= E\{\log_2 \tfrac{1}{p(x,y)}\} \end{aligned}$$

(a) Halle una expresión que relacione ambas definiciones y dé una interpretación a dicha relación.

(b) Defina la información mutua $I\{X;Y\}$ a partir de la expresión anterior y dé una interpretación a dicha definición.

★ Ejercicio 3

Sea $p(x,y)$ dado por las siguientes probabilidaes hallar:

$X \backslash Y$	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	0	$\frac{1}{3}$

(a) $H\{X\}$ y $H\{Y\}$.

(b) $H\{X|Y\}$ y $H\{Y|X\}$.

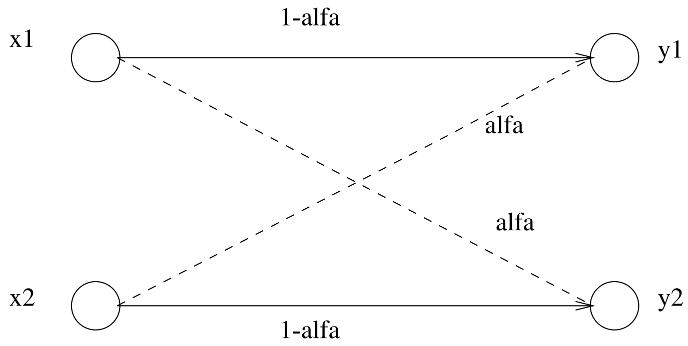
(c) $H\{X,Y\}$.

(d) $I\{X;Y\}$.

(e) Realice un diagrama de Venn relacionando todas las cantidades anteriores.

⋆ Ejercicio 4

Canal Binario Simétrico. Considerar el modelo de un canal binario simétrico, representado en la figura:



donde: $P(x_1) = p$ y $P(y_1|x_2) = P(y_2|x_1) = \alpha$.

(a) Calcular la información mutua en función de p y α

(b) Discutir qué valores toma α cuando la potencia del ruido es muy pequeña o muy grande.

✶ Ejercicio 5

Se considera un archivo en una computadora de largo M bits (con M suficientemente grande a todos los efectos), tal que la ocurrencia de 0s y 1s son iid y con probabilidad de que un bit sea 0 igual a p_0. Dado que p_0 es relativamente alto, se considera una codificación alternativa donde se indicará cuántos 0s sucesivos hay previo a cada 1 (una variación del denominado *Run-length encoding*, donde supondremos que todos los archivos terminan en 1). Por ejemplo, si el archivo comienza con 00110001, las primeras 3 salidas del codificador serán $X_i = 2, 0, 3$ para $i = 1, 2, 3$.[1]

(a) Calcule el largo medio del archivo re-codificado en función de p_0 y M. Grafíquelo en función de p_0 junto con el largo promedio que obtendría un codificador binario óptimo y verifique que este último es mayor para muchos valores de p_0. Argumente por qué esto no contradice el teorema de codificación de fuente.

(b) Calcule la entropía del archivo re-codificado. Verifique que es igual a la del archivo original. Argumente por qué es esperable este resultado.

[1] Pueden ser útiles para el ejercicio las igualdades $\sum_{k\geq 0} kp^{k-1} = 1/(1-p)^2$ y $\sum_{k\geq 0} p^k = 1/(1-p)$.

Solución

Ejercicio 1

Debemos probar que si X e Y son símbolos discretos, aleatorios e independientes, se cumple que:

$$H\{X;Y\} = H\{X\} + H\{Y\}$$

Partimos entonces de la definición:

$$\begin{aligned} H\{X;Y\} &= \textstyle\sum_{i,j} p(x_i, y_j) \log_2 \frac{1}{p(x_i,y_j)} \\ &= \textstyle\sum_{i,j} p(x_i)p(y_j) \left(\log_2 \frac{1}{p(x_i)} + \log_2 \frac{1}{p(y_i)}\right) \\ &= H\{X\} + H\{Y\} \end{aligned}$$

Ejercicio 2

(a)

$$H(X,Y) = E\{\log_2 \frac{1}{p(x,y)}\}$$

$$= \sum_{X,Y} p(x,y).log_2 \left(\frac{1}{p(x,y)}\right)$$

$$= \sum_{X,Y} p(x,y).log_2 \left(\frac{1}{p(y)p(x|y)}\right)$$

$$= \sum_{X,Y} p(x,y). \left(log_2 \left(\frac{1}{p(y)}\right) + log_2 \left(\frac{1}{p(x|y)}\right)\right)$$

$$= \sum_{X,Y} p(x,y).log_2 \left(\frac{1}{p(y)}\right) + \sum_{X,Y} p(x,y).log_2 \left(\frac{1}{p(x|y)}\right)$$

$$= E\{\log_2 \frac{1}{p(y)}\} \ + E\{\log_2 \frac{1}{p(x|y)}\}$$

$$= H(Y) + H(X|Y)$$

De la expresión anterior, al ser $H(X,Y)$ la incertidumbre de las variables conjuntas X e Y; se puede deducir que $H(X|Y)$ es la incertidumbre faltante de la variable X tras conocer la variable Y.

(b) Retomando la expresión de la parte anterior se deduce que

$$H(X,Y) = H(Y) + H(X|Y) = H(X) + H(Y|X),$$

y por lo tanto

$$H(X,Y) - H(X|Y) - H(Y|X) = H(Y) - H(Y|X) = H(X) - H(X|Y) = I(X,Y).$$

Definimos la información como la resta entre la entropía conjunta (la cuál representa la máxima cantidad de incertidumbre) y las entropías condicionales de ambas variables. El resultado entonces es la reducción de incertidumbre de una variable aleatoría al conocer la otra, o en otras palabras, la información que me aporta conocer conocer la variable Y respecto de X (o viceversa).

Ejercicio 3

(a)

$$H\{X\} = -\sum_{x\in X} p_X(x) log_2 p_X(x) = -\frac{2}{3}log_2\left(\frac{2}{3}\right) - \frac{1}{3}log_2\left(\frac{1}{3}\right) \approx 0.918 \text{ bits.}$$

$$H\{Y\} = -\sum_{y\in Y} p_Y(y) log_2 p_Y(y) = -\frac{2}{3}log_2\left(\frac{2}{3}\right) - \frac{1}{3}log_2\left(\frac{1}{3}\right) \approx 0.918 \text{ bits.}$$

(b)

$$H\{X|Y\} = \sum_{y\in Y} H\{X|Y=y\} p_Y(y) = \frac{2}{3}\Omega\left(\frac{1}{2}\right) + \frac{1}{3}\Omega(0) \approx 0.67 \text{ bits.}$$

$$H\{Y|X\} = \sum_{x\in X} H\{Y|X=x\} p_X(x) = \frac{2}{3}\Omega\left(\frac{1}{2}\right) + \frac{1}{3}\Omega(0) \approx 0.67 \text{ bits.}$$

(c)

$$H\{X,Y\} = -\sum_{x,y} p_{X,Y}(x,y) log_2(x,y) = -3 \times \left[\frac{1}{3}log_2\left(\frac{1}{3}\right)\right] \approx 1.585 \text{ bits.}$$

(d) Por definición:

$$I\{X;Y\} = H\{X,Y\} - H\{X|Y\} - H\{Y|X\} \approx 0.251 \text{ bits.}$$

(e) Ver figura[2] 1.

Ejercicio 4

(a)

$$I(X,Y) = \sum_{X,Y} P(x_i,y_j).log_2\frac{P(x_i,y_j)}{P(x_i)P(y_j)} = \sum_{X,Y} P(x_i,y_j).log_2\frac{P(y_j/x_i)}{P(y_j)}$$

$$P(y_1) = P(x_1)P(y_1/x_1) + P(x_2)P(y_1/x_2) = p(1-\alpha) + (1-p)\alpha$$
$$P(y_2) = P(x_1)P(y_2/x_1) + P(x_2)P(y_2/x_2) = p\alpha + (1-p)(1-\alpha)$$

$$I(X,Y) = P(x_1)P(y_1/x_1).log_2\frac{P(y_1/x_1)}{P(y_1)} + P(x_2)P(y_1/x_2).log_2\frac{P(y_1/x_2)}{P(y_1)} + ...$$

$$...P(x_1)P(y_2/x_1).log_2\frac{P(y_2/x_1)}{P(y_2)} + P(x_2)P(y_2/x_2).log_2\frac{P(y_2/x_2)}{P(y_2)}$$

Sustituyendo queda:

[2] By KonradVoelkel - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=11245361. 28 de mayo de 2021.

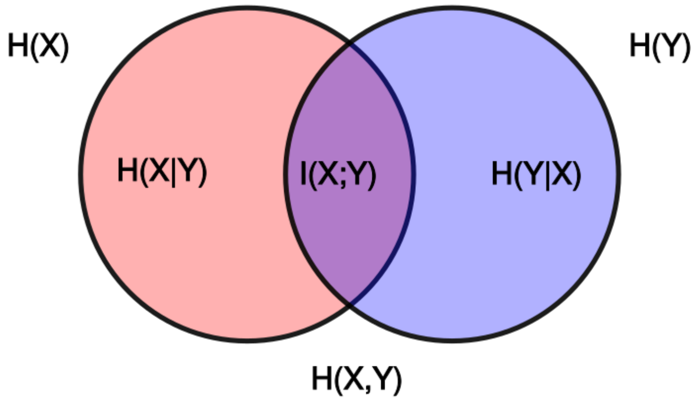


Figura 1: Diagama de Venn conceptual relacionando entropía, entropía conjunta, entropía condicional e información mutua.

$$I(p,\alpha) = p(1-\alpha).log_2\frac{(1-\alpha)}{p(1-\alpha)+(1-p)\alpha} + (1-p)\alpha.log_2\frac{\alpha}{p(1-\alpha)+(1-p)\alpha} + ...$$

$$...p\alpha.log_2\frac{\alpha}{p\alpha+(1-p)(1-\alpha)} + (1-p)(1-\alpha).log_2\frac{1-\alpha}{p\alpha+(1-p)(1-\alpha)}$$

(b) Si la potencia del ruido es muy grande, se tiene igual probabilidad de recibir y_1 o y_2 (el ruido hace que lo recibido sea aleatorio e independiente de la entrada). Por lo que en este caso α toma valores cercanos a $\frac{1}{2}$.
Si la potencia del ruido es muy pequeña, entonces cuando se envía un símbolo (por ej x_1), siempre se recibe el mismo símbolo en la recepción ya sea y_1 o y_2. En este caso α toma un valor cercano a 0 o a 1.

Ejercicio 5

(a) El largo del archivo re-codificado será cuántos 1s hay en el archivo original. Por lo tanto, y dado que la probabilidad de cada 1 es $(1-p_0)$, resulta $M(1-p_0)$. Por otro lado, un codificador binario óptimo va a lograr como mínimo un largo medio igual a la entropía de cada bit multiplicado por M (es decir, la entropía del archivo original). Por lo tanto, el tamaño mínimo será:

$$\overline{L}_{\text{mín}} = -M\left(p_0 \log_2(p_0) + (1-p_0)\log_2(1-p_0)\right). \quad (1)$$

Es fácil verificar realizando la curva de la entropía y $(1-p_0)$ que para todos los valores de p_0 de interés de este ejercicio (p_0 cercano a 1) la entropía es mayor (figura 2). Incluso para el caso de $p_0 = 0.5$ el largo medio del código binario óptimo es M, pero para el código estudiado aquí es $M/2$. La aparente contradicción con el teorema de codificación es que el código estudiado aquí no es binario y usa todos los naturales.

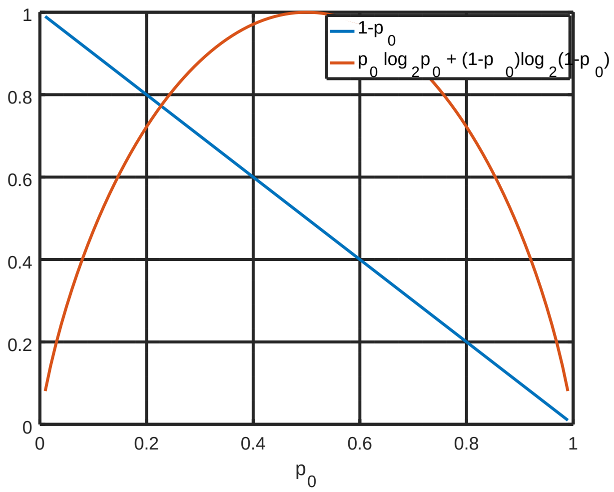


Figura 2: La entropía de cada bit y la función $(1-p_0)$.

(b) Como cada bit es iid, la cantidad de 0s consecutivos antes de cada 1 también serán iid. En particular, la probabilidad de que la cantidad de veces seguidas que aparece un 0 sea k es:

$$P(X=k) = (p_0)^k(1-p_0).$$

Tenemos entonces que los símbolos posibles en nuestro alfabeto son los naturales, son iid y con la distribución anterior. Por lo tanto, la entropía total del archivo re-codificado será su largo medio por la entropía de cada símbolo.
Esta entropía es:

$$\begin{aligned} H\{X\} &= -\sum_{k=0}^{\infty} p_0^k(1-p_0)\log_2\left(p_0^k(1-p_0)\right) \\ &= -\sum_k p_0^k(1-p_0)\log_2\left(p_0^k\right) - \sum_k p_0^k(1-p_0)\log_2\left(1-p_0\right) \end{aligned} \tag{2}$$

$$\Rightarrow H\{X\} = -(1-p_0)\log_2(p_0)\sum_k p_0^k k - (1-p_0)\log_2(1-p_0)\sum_k p_0^k, \tag{3}$$

donde se usaron propiedades básicas de logaritmos. Usando las fórmulas para series que se dan en la letra lo anterior resulta:

$$H\{X\} = -\log_2(p_0)\frac{p_0}{1-p_0} - \log_2(1-p_0). \tag{4}$$

Como el largo medio del archivo re-codificado es $(1-p_0)M$, la entropía total resulta

$$\begin{aligned} M(1-p_0)H\{X\} &= M(1-p_0)\left(-\log_2(p_0)\tfrac{p_0}{1-p_0} - \log_2(1-p_0)\right) \\ &= -M\left(p_0\log_2(p_0) + (1-p_0)\log_2(1-p_0)\right), \end{aligned} \tag{5}$$

igual a la del archivo original. Esto debe ser así pues simplemente estamos sustituyendo una representación del mismo archivo por otra, sin perder información en el proceso (es decir, se puede recuperar el archivo original con el re-codificado). En la entropía entran únicamente las probabilidades de cada evento o símbolo (o archivo en este caos), y no su representación.