Introducción a la Teoría de la Información
Primer parcial

1 de abril de 2024

Escribir con lapicera o lápiz con buen contraste, que se lea bien. Las soluciones no son sólo cuentas; desarrollar en forma prolija y explicar lo que van haciendo.

Problema 1 (5 puntos)

Sea $X \in \mathcal{X}$, $X \sim p$, una variable aleatoria, donde $\mathcal{X}$ es un conjunto finito. Complete la siguiente demostración de una cota superior para $H(X)$ y condiciones de igualdad.

1. Sea q ...
2. $D(p||q) = \log |\mathcal{X}| - H(X)$
3. Entonces $H(X) \leq \ldots$, con igualdad si y solo si ....

Solución:

Ver teorema 2.6.4 en Cover.

Problema 2 (5 puntos)

Sean X, Y variables aleatorias discretas.

1. Mostrar que $H(X) - H(X|Y) = H(Y) - H(Y|X)$.

   (Estas diferencias entre entropía y entropía condicionada, que son iguales entre sí, dan como resultado la información mutua entre X e Y, $I(X;Y)$).

Sea ahora X una variable aleatoria en un conjunto $\mathcal{X}$ finito, con valores posibles x_i a cuyas probabilidades llamaremos p_i. Sea $\mathcal{S}$ un subconjunto de $\mathcal{X}$.

Se define una variable aleatoria Y como sigue:
$Y = 1$ si $X = x_i \in \mathcal{S}$
$Y = 0$ si $X = x_i \notin \mathcal{S}$

2. $\alpha = P(Y = 1)$. Expresar α en función de las probabilidades p_i.

3. Expresar $H(Y)$ en función de α.
4. Hallar $I(X;Y)$ en función de α a partir de $H(Y)$ y $H(Y|X)$.
5. Hallar nuevamente $I(X;Y)$, a partir de $H(X)$ y $H(X|Y)$.
6. Relacionar la diferencia $H(X) - H(X|Y)$ con el juego de adivinar un elemento de X mediante preguntas de respuesta binaria (sí o no). ¿Qué valor de α resulta más conveniente?

Solución:

1. Descomponiendo $H(X,Y)$ de dos formas distintas usando la regla de la cadena obtenemos

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

y reordenando los términos llegamos a

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

2. Tenemos $\alpha = \sum_{x_i \in \mathcal{S}}(p_i)$. Además se cumple $1-\alpha = \sum_{x_i \notin \mathcal{S}}(p_i)$.
3. $H(Y) = H(\alpha)$, donde $H(\alpha)$ denota la función de entropía binaria, $H(\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$.
4. $I(X;Y) = H(Y) - H(Y|X) = H(\alpha)$, pues $H(Y|X) = 0$ porque Y es función de X.
5. En la expresión $I(X;Y) = H(X) - H(X|Y)$ calculamos primero $H(X|Y)$. Tenemos

$$\begin{aligned} H(X|Y) &= P(Y=0)H(X|Y=0) + P(Y=1)H(X|Y=1) && (1)\\ &= (1-\alpha)H(X|Y=0) + \alpha H(X|Y=1)\,. && (2) \end{aligned}$$

Para todo $x_i \in \mathcal{S}$ tenemos

$$P(X = x_i|Y=1) = \frac{P(X=x_i, Y=1)}{P(Y=1)} = \frac{P(X=x_i)}{P(Y=1)} = \frac{p_i}{\alpha}\,,$$

y para todo $x_i \in \mathcal{X} \setminus \mathcal{S}$ tenemos $P(X=x_i|Y=1) = 0$.
Por lo tanto, para el término $\alpha H(X|Y=1)$ de (2) podemos escribir

$$\begin{aligned} \alpha H(X|Y=1) &= -\alpha \sum_{x_i\in\mathcal{S}} \frac{p_i}{\alpha}\log\frac{p_i}{\alpha} && (3)\\ &= -\sum_{x_i\in\mathcal{S}} p_i \log p_i + \sum_{x_i\in\mathcal{S}} p_i \log \alpha && (4)\\ &= \left(-\sum_{x_i\in\mathcal{S}} p_i\log p_i\right) + \alpha\log\alpha\,. && (5) \end{aligned}$$

Análogamente, para el término $(1-\alpha)H(X|Y=0)$ de (2) obtenemos

$$(1-\alpha)H(X|Y=0) = \left(-\sum_{x_i \in \mathcal{X}\setminus\mathcal{S}} p_i \log p_i \right) + (1-\alpha)\log(1-\alpha)\,. \qquad (6)$$

Reemplazando (5) y (6) en (2) llegamos a

$$H(X|Y) = \left(-\sum_{x_i \in \mathcal{X}} p_i \log p_i \right) + \alpha\log\alpha + (1-\alpha)\log(1-\alpha) = H(X) - H(\alpha)\,.$$

Recordando que $I(X;Y) = H(X) - H(X|Y)$, concluimos que

$$I(X;Y) = H(\alpha)\,.$$

6. La diferencia $H(X) - H(X|Y)$ representa la disminución media de incertidumbre con respecto al elemento que se quiere adivinar, X, al obtener la respuesta a la presunta binaria ¿pertenece X a $\mathcal{S}$? Esta reducción será máxima (se obtiene más información) cuando se maximiza $H(\alpha)$, es decir, cuando se elige $\mathcal{S}$ tal que $\alpha = \frac{1}{2}$ o lo más cercano posible.

Problema 3 (5 puntos)

Sea $\{X_n\}_{n\geq 1}$ un proceso de Markov con conjunto de estados $\mathcal{X} = \{0,1,2\}$, donde X_1 tiene distribución uniforme en $\mathcal{X}$ y la matriz de probabilidades de transición es

$$P = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$$

Considerar el proceso $\{Z_n\}_{n\geq 1}$ definido por

$$\begin{aligned} Z_1 &= X_1 && (7) \\ Z_i &= X_i - X_{i-1} \pmod 3,\ i > 1\,. && (8) \end{aligned}$$

1. El proceso $\{X_n\}_{n\geq 1}$, ¿es estacionario? Justificar.
2. Calcular una tasa de entropía de $\{Z_n\}_{n\geq 1}$.

Solución:

1. El proceso es estacionario porque $(1/3, 1/3, 1/3)P = (1/3, 1/3, 1/3)$.
2. Como $Z_1 \dots Z_n$ es una función biyectiva de $X_1 \dots X_n$, tenemos $H(Z_1 \dots Z_n) = H(X_1 \dots X_n)$ y por lo tanto

$$\lim_{n\to\infty} \frac{H(Z_1 \dots Z_n)}{n} = \lim_{n\to\infty} \frac{H(X_1 \dots X_n)}{n} = H(X_2|X_1) = 3/2 \text{ bits}\,.$$
