Introducción a la Teoría de la Información
Segundo parcial

24 de abril de 2024

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

Problema 1 (5 puntos)

Considere este enunciado: Sea C un código D-ario instantáneo para una variable aleatoria $X \sim p$. Se cumple $L(C) \geq H_D(X)$, y se da la igualdad si y solo si $p_i = D^{-l_i}$ para todo i.

Complete una demostración del enunciado justificando los siguientes pasos.

1. Para cada i sea $q_i = D^{-l_i}/K$, donde $K = \ldots$
2. $D(p||q) = -H_D(X) + L(C) + \log_D K$.
3. Si $p_i = D^{-l_i}$ para todo i, entonces $L(C) = H_D(X)$.
4. Recíprocamente, si $L(C) = H_D(X)$, entonces $p_i = D^{-l_i}$ para todo i.

Solución:

1. Para cada i sea $q_i = D^{-l_i}/K$, donde $K = \sum_i D^{-l_i}$.
2. Operando a partir de la definición de divergencia tenemos

$$\begin{aligned}
D(p||q) &= \sum_i p_i \log_D \frac{p_i}{q_i} = \sum_i p_i \log_D p_i - \sum_i p_i \log_D q_i \\
&= -H_D(X) - \sum_i p_i \log_D \frac{D^{-l_i}}{K} \\
&= -H_D(X) - \sum_i p_i \log_D D^{-l_i} + \sum_i p_i \log_D K \\
&= -H_D(X) + \sum_i p_i l_i + \log_D K \\
&= -H_D(X) + L(C) + \log_D K\,.
\end{aligned}$$

Por lo tanto se cumple

$$L(C) - H_D(X) = D(p||q) - \log_D K \geq 0\,, \tag{1}$$

donde la desigualdad surge de que la divergencia es no negativa y, como por la desigualdad de Kraft tenemos $K \leq 1$, el término $-\log_D K$ también es no negativo.

3. Si $p_i = D^{-l_i}$ para todo i, entonces $K = \sum_i p_i = 1$, por lo cual $q_i = p_i$ para todo i. Esto implica que los dos términos de la derecha en (1) son 0 y en consecuencia $L(C) = H_D(X)$.
4. Recíprocamente, si $L(C) = H_D(X)$, entonces cada término de la derecha en (1) debe ser 0, ya que ambos son no negativos. El hecho de que $-\log_D K = 0$ implica que $K = 1$ y por lo tanto $q_i = D^{-l_i}$ para todo i. Por otra parte el hecho de que $D(p||q) = 0$ implica que $p_i = q_i$ para todo i, con lo cual concluimos que $p_i = D^{-l_i}$ para todo i.

Problema 2 (5 puntos)

Se te dan 6 botellas de vino, $b_1, b_2, \ldots, b_6$, de las cuales exactamente una está en mal estado. A partir de la inspección de las botellas se determina que la probabilidad p_i de que la i-ésima botella esté en mal estado está dada por $(p_1, p_2, ..., p_6) = \left(\frac{8}{23}, \frac{6}{23}, \frac{4}{23}, \frac{2}{23}, \frac{2}{23}, \frac{1}{23}\right)$.

Para determinar el vino malo se realizan mezclas de vinos, probando cada mezcla para confirmar o descartar la presencia del vino malo en la mezcla, y procediendo con diferentes mezclas hasta determinar cuál es el vino malo.

1. ¿Cuál es el mínimo número esperado de pruebas requeridas para determinar el vino malo?

2. Supongamos que el vino malo es b_2. Siguiendo una estrategia que minimiza la cantidad esperada de pruebas, ¿qué secuencia de pruebas de mezclas se realiza hasta encontrar el vino malo?

3. ¿Cuál es la secuencia si el vino malo es b_4?

Solución:

1. Sea $X \in \{b_1, b_2, \ldots, b_6\}$ una variable aleatoria que representa al vino malo. Un posible código de Huffman para X asigna las palabras de código $(00, 01, 10, 110, 1110, 1111)$. El largo esperado de código es

$$L = 2 \times \left(\frac{8}{23} + \frac{6}{23} + \frac{4}{23}\right) + 3 \times \frac{2}{23} + 4 \times \left(\frac{2}{23} + \frac{1}{23}\right) = 2,35\,.$$

Partiendo de un conjunto de candidatos S inicialmente igual a $\{b_1, b_2, \ldots, b_6\}$, realizamos pruebas de mezclas secuencialmente, descartando elementos de S en cada paso, hasta llegar a un conjunto S con un único elemento, que es necesariamente el vino malo.

Un algoritmo concreto para identificar el vino malo es el siguiente:

1. Hacer $S = \{b_1, b_2, \ldots, b_6\}$, $i = 1$.
2. Mezclar los vinos del conjunto M formado por todos los vinos de S cuya palabra de código tiene 0 en la posición i.
3. Probar la mezcla M. Si sabe mal, hacer $S = M$; en caso contrario hacer $S = S \setminus M$.
4. Si S tiene un solo elemento, b, terminar: b es el vino malo. En caso contrario incrementar i y pasar al paso 2..

2. Probamos la mezcla $\{b_1, b_2\}$. Como sabe mal, probamos b_1, que sabe bien y por lo tanto deducimos que b_2 es el vino malo.
3. Probamos la mezcla $\{b_1, b_2\}$. Como sabe bien, deducimos que el vino malo está en $S = \{b_3, \ldots, b_6\}$. Probamos b_3, que sabe bien y por lo tanto deducimos que el malo está en $S = \{b_4, \ldots, b_6\}$. Probamos b_4 y lo identificamos como el vino malo.

Problema 3 (5 puntos)

Sea X una variable aleatoria sobre un alfabeto de cuatro símbolos, $\mathcal{X} = \{a, b, c, d\}$, con distribución $p = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{12}\right)$.

1. Muestre que hay códigos de Huffman para X con largos de código $(1, 2, 3, 3)$ y $(2, 2, 2, 2)$.
2. Concluya que existen códigos óptimos tales que a alguno de los símbolos del alfabeto le asignan una palabra de código más larga que la asignada por un código de Shannon.

Solución:

1. Al juntar los símbolos c y d, que tienen probabilidades $1/4$ y $1/12$, en el algoritmo de Huffman obtenemos un nuevo símbolo, llamémosle e, con probabilidad $1/3$. A continuación se prosigue recursivamente con el alfabeto $\{a, b, e\}$ con distribución uniforme $(1/3, 1/3, 1/3)$, que tiene más de una solución dependiendo de cómo se desempate para realizar la siguiente agrupación de símbolos. Una solución asigna largos $(1, 2, 2)$, que al desagrupar el símbolo e nos da largos $(1, 2, 3, 3)$ para X. Otra solución posible asigna largos $(2, 2, 1)$, que al desagrupar el símbolo e nos da largos $(2, 2, 2, 2)$ para X.

2. El símbolo c recibe una palabra de código de largo $\lceil -\log \frac{1}{4} \rceil = 2$ en un código de Shannon. Por otra parte, como vimos en la parte anterior, existe un código de Huffman (que es óptimo), que asigna una palabra de código de largo 3 al símbolo c.
