Compresión de datos sin pérdida
Práctico 2

Ejercicio 1 *Típos de código*
Considere un código con palabras de código $\{0, 01\}$.

1. ¿Es no singular?
2. ¿Es instantáneo?
3. ¿Es unívocamente decodificable?

Ejercicio 2 *Códigos malos*
¿Cuál(es) de los siguientes códigos no pueden ser óptimos para ninguna distribución de probabilidad?

1. {0, 10, 11}
2. {00, 01, 10, 110}
3. {01, 10}

Ejercicio 3 *Propiedades de códigos óptimos*
Sea X una variable aleatoria sobre un alfabeto finito $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$, con distribución de probabilidad determinada por $p = [p_1, p_2, \ldots, p_m]$, y sea C un código de prefijo óptimo para X, con largos de palabras de código $[\ell_1, \ell_2, \ldots, \ell_m]$. Mostrar que

1. si $p_i > p_j$, entonces $\ell_i \leq \ell_j$,
2. para toda palabra de código de largo máximo existe otra del mismo largo que sólo difiere en el último símbolo.

Supongamos que $p_{m-1} \leq p_i$ y $p_m \leq p_i$ para todo i, $1 \leq i < m-1$. Mostrar que existe un código de prefijo óptimo, C', tal que las palabras de código asociadas a a_m y a_{m-1} difieren entre sí sólo en el último símbolo.

Ejercicio 4 *Huffman*
Consideramos dos variables aleatorias, $X \sim p$, $Y \sim q$. El alfabeto de X tiene seis símbolos, con probabilidades

$$p = (0{,}3,\ 0{,}2,\ 0{,}15,\ 0{,}15,\ 0{,}1,\ 0{,}1)\,.$$

El alfabeto de Y tiene siete símbolos, con probabilidades

$$q = (0{,}3,\ 0{,}25,\ 0{,}15,\ 0{,}1,\ 0{,}1,\ 0{,}05,\ 0{,}05)\,.$$

Construir códigos de Huffman para X y para Y, y hallar en cada caso el largo de código medio.

Ejercicio 5 *Códigos óptimos para distribuciones uniformes*
Considere una variable aleatoria X que toma valores en un alfabeto con m elementos, $\mathcal{A} = \{a_1, a_2, \ldots a_m\}$, con distribución uniforme.

1. Mostrar que en todo código instantáneo óptimo para X los largos de las palabras de código difieren en no más de 1.

2. Calcule el largo de código medio óptimo, $L(X)$.

3. ¿Para qué valores de m se cumple que $L(X)$ es igual a $H(X)$? Justifique.

Ejercicio 6
Mostrar que si una variable aleatoria toma un determinado valor con probabilidad mayor que 0.4, entonces todo código óptimo para esa variable aleatoria tiene al menos una palabra de largo 1.

Ejercicio 7 *Huffman y Shannon*
Sea X una variable aleatoria sobre un alfabeto de cuatro símbolos con distribución $p = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{12}\right)$.

1. Mostrar que hay códigos de Huffman para X con largos de código $(1, 2, 3, 3)$ y $(2, 2, 2, 2)$.

2. Concluir 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.

Ejercicio 8 *Kraft*
Un código libre de sufijos tiene la propiedad de que ninguna palabra de código es sufijo de otra.

1. Mostrar que un código libre de sufijos es unívocamente decodificable.

2. Sea $\{\ell_1, \ell_2, \ldots, \ell_m\}$ un conjunto de enteros positivos. Mostrar que si

$$\sum_{i=1}^{m} 2^{-\ell_i} \leq \frac{1}{2},$$

entonces existe un código con palabras de largo $\ell_1, \ell_2, \ldots, \ell_m$ que es simultáneamente libre de prefijos y de sufijos.