Compresión de datos sin pérdida
Tercera prueba escrita

8 de julio de 2021

Ejercicio 1 *(30 puntos)*
Consideremos el codificador aritmético estudiado en el curso, con precisión numérica de K dígitos en base D para la representación numérica de los anchos de intervalo T_i, $0 \leq i \leq n$. Recordemos que para todo i, $1 \leq i \leq n$, se cumple (no es necesario probarlo)

$$T_{i-1} Q_i[x_i] \Big(1 - D^{1-K} \Big) \leq T_i \,. \tag{1}$$

Mostrar que para toda secuencia, x^n, el largo de código $L(x^n)$ satisface

$$\frac{L(x^n)}{n} \leq -\frac{\log_D Q(x^n)}{n} + \frac{2}{n} + \nu_D(K) \,, \tag{2}$$

donde $\nu_D(K) = -\log_D \Big(1 - D^{1-K} \Big)$ y $Q(x^n) = \prod_{i=1}^{n} Q_i[x_i]$.
Sugerencia: Tomar logaritmos en (1) y sumar en i. Recordar que

$$L(x^n) = \lceil -\log_D T_n \rceil + 1 \,.$$

Ejercicio 2 *(35 puntos)*
Sea $Q_n^{\text{L}}(x^n)$ la probabilidad asignada a una secuencia binaria x^n por la asignación secuencial de Laplace, y sea $P_{ML}(x^n)$ la máxima verosimilitud de x^n con respecto a la familia de modelos de Bernoulli, $\mathcal{C}_B$.

1. Calcular $Q_n^{\text{L}}(x^n)$ para $x^n = 001110$.
2. Mostrar que $Q_n^{\text{L}}(x^n) \leq P_{ML}(x^n)$ para toda secuencia binaria x^n.

   Sugerencia: Recordar la definición de Q_n^{L} como mezcla de distribuciones sobre $\mathcal{X}^n$.

Ejercicio 3 *(35 puntos)*
Considere un código de Golomb PO2, G_k^*, de parámetro k. Denotamos con $|G_k^*(x)|$ al largo de código que G_k^* asigna al entero x.

1. Demostrar que $|G_k^*(x)| \geq k + 1$ para todo x.
2. Demostrar que para todo $j \geq 0$, el conjunto

   $$S_j = \Big\{ x \in \mathbb{N} \;\Big|\; |G_k^*(x)| = k + 1 + j \Big\}$$

   contiene exactamente 2^k enteros. ¿Cuáles son estos enteros?