Reading Assignment

- Read Chapter 5 of the Cover and Thomas book (first 10 sections) and Theorem 4.1 of the Csiszár and Körner book.

Homework Questions

Q1 Consider a source $X$ with pmf

$$P(0) = \frac{1}{2}, \quad P(1) = P(2) = \frac{1}{8}, \quad P(3) = P(4) = P(5) = P(6) = \frac{1}{16}.$$ 

a) Find $L_0(X), \overline{L}_0(X), L_{\text{unique}}(X),$ and $L_{\text{prefix}}(X)$.

b) Find $L_{0.25}(X)$ and $\overline{L}_{0.25}(X)$.

You can assume that the empty sequence $\emptyset$ is a valid codeword.

Q2 a) Show that $L_{\text{prefix}}(X) = H(X)$ if and only if $-\log P(x)$ is an integer for every $x$.

b) Exhibit an example where $L_{\text{prefix}}(X) > H(X)$. Compute the exact value of $L_{\text{prefix}}(X)$ for your example.

Q3 Consider the source $X$ with pmf

$$P(1) = \frac{1}{3}, \quad P(2) = \frac{1}{6}, \quad P(3) = \frac{1}{6}, \quad P(4) = \frac{1}{3}.$$ 

a) Find the Huffman code, Shannon-Fano code, and Elias code for this source.

b) Which of these constructions are unique?

c) Consider the sequence of symbols 3142 produced by a DMS with a common distribution $P$. Determine the output of the arithmetic code (discussed in the class) applied to this sequence. Apply the corresponding decoding algorithm to the compressed sequence and verify the FIFO property.

Q4 For a DMS $X^n$ with a (unknown) pmf $P$, consider the following compression scheme: For each sequence $x$ store its type $P_x$ and its enumeration in the type class $T_{P_x}$.

a) Is this code nonsingular? Is it uniquely decodeable? Is it prefix-free?

b) For a pmf $P$, denote by $\overline{L}_n(P)$ the average length of this code for a pmf $P$. Determine an upper bound for the gap to optimality

$$\sup_P \overline{L}_n(P) - nH(P).$$

Conclude that for every pmf $P$ the scheme above attains the optimal rate $\overline{R}_0$. 