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

1. 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 $L(X)$ and $L^u(X)$ and give codes which achieve each of these quantities.
b) Find $L_{0.25}(X)$ and a code which achieves it.
c) Can you find the exact value of $L_{0.25}(X)$? If so, what is it? If not, why not?
You can assume that the empty sequence $\emptyset$ is a valid codeword.

2. Exhibit an example where $L^p(X) > H(X)$. Compute the exact value of $L^p(X)$ for your example.

3. 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 a Huffman code, a Shannon-Fano code, and a Shannon-Fano-Elias code for this source.
b) Which of these constructions are unique?

4. For the pmf $P$ of the previous question, consider the sequence of symbols 3142 produced by a DMS with a common distribution $P$. Determine the output of the (infinite precision) arithmetic code applied to this sequence. Apply the corresponding decoding algorithm to the compressed sequence and verify the FIFO property.

5. Given random variables $X$ and $Y$ taking values in finite sets $\mathcal{X}$ and $\mathcal{Y}$, respectively, and with a joint pmf $P_{XY}$, consider the following extension of variable-length source codes:

The encoder observes $X$ and $Y$ and compresses $X$ as $Z = e(X, Y) \in \{0, 1\}^*$. The decoder observes $Y$ and the output $Z$ of the encoder and forms an estimate $\hat{X} = d(Y, Z)$. For every $(x, y) \in \mathcal{X} \times \mathcal{Y}$, denote by $l(x, y)$ the length of the bit string $e(x, y)$. The average length of this code is given by

\[ L_e(X|Y) = \sum_{x,y} P_{XY}(x,y) l(x,y). \]

A code is prefix-free if for every $y \in \mathcal{Y}$ the codeword set $\{e(x, y) : x \in \mathcal{X}\}$ is prefix-free. Denote by $\overline{L}^p(X|Y)$ the minimum average length of a prefix-free code.
a) Show that
\[ \overline{I}_p(X|Y) \leq \overline{I}_p(X). \]

b) Is the following statement true or false? If you say true, prove it. If you say false, provide a counter example.
\[ \overline{I}_p(X|Y) = \overline{I}_p(X) \text{ if and only if } X \text{ and } Y \text{ are independent.} \]

6. Let \( X_1, \ldots, X_n \) be i.i.d. with a common (unknown) pmf \( P \) over a finite alphabet \( \mathcal{X} \). Show that the maximum likelihood estimate of \( P \) upon observing \( X^n = x \) is the type \( P_x \) of \( x \).

7. Let \( X_1, \ldots, X_n \) be i.i.d. with a common (unknown) pmf \( P \) over a finite alphabet \( \mathcal{X} \). What is the conditional distribution of \( X^n \) given the event \( X^n \in \mathcal{T}_Q \) for some type \( Q \in \mathcal{T} \)?

8. Let \( X_1, \ldots, X_n \) be i.i.d. with a common (unknown) pmf \( P \) over a finite alphabet \( \mathcal{X} \). Let \( Q \in \mathcal{T} \) denote the type of the random sequence \( X^n \). Show that for every \( \epsilon > 0 \)
\[ P(D(Q||P) \geq \epsilon) \leq (n+1)^{|X|}2^{-n\epsilon}. \]

Remark: This result can be used to show that the type of the random sequence converges to \( P \) almost surely.

9. Show that
\[ \frac{1}{n+1} \cdot 2^{nh(x)} \leq \binom{n}{k} \leq 2^{nh(x)}, \]
where \( h \) denotes the binary entropy function.

10. For \( \delta \in [0, 1/2] \), denote by \( V_{n,\delta} \) the “volume of the \( n \)-dimensional Hamming sphere of radius \( n\delta \),” i.e., the number of sequences in \( \{0,1\}^n \) with less than or equal to \( n\delta \) 1s. Determine the limit
\[ \lim_{n \to \infty} \frac{1}{n} \log V_{n,\delta}. \]

Hint: What is the notion of type for the binary alphabet? This hint applies to the previous question as well.

11. Let \( \mathcal{X} \) be a finite alphabet. For a sequence \( x \in \mathcal{X}^n \), the profile \( N_x \) of the sequence \( x \) is given by the vector \( N_x = (N_1, \ldots, N_n) \) where \( N_i \) denotes the number of symbols that have appeared \( i \) times. For instance, the profile of the sequence \( abcadabca \) is \( (1,2,1,0,\ldots,0) \). Denote by \( \mathcal{S}_N \) the set of sequences in \( \mathcal{X}^n \) with a fixed profile \( N \).

Show that the following fixed-length source code is universally rate optimal:

Given a fixed rate \( R > 0 \), let \( A \) denote the set of sequences \( x \) such that the profile \( N_x = (N_0, \ldots, N_n) \) satisfies
\[ R > \sum_{i=0}^{n} N_i \cdot \binom{i}{n} \log \left( \frac{n}{i} \right). \]

The source code is given by simply mapping the sequences in \( A \) to binary sequences of a fixed length \( nR \) (check that this can be done). All the sequences outside \( A \) are mapped to the all 0 sequence of the same length. The decoder simply returns the sequence in \( A \) corresponding to the stored binary sequence.