Information Theory I

Fall Semester 2017

Prof. Dr. A. Lapidoth

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Exercise 5 of October 18, 2017

http://www.isi.ee.ethz.ch/teaching/courses/it1.html

Problem 1 — *Huffman Coding*

Consider the chance variable X taking on the values $(x_1, x_2, \ldots, x_6)$ with respective probabilities $(0.49, 0.17, 0.15, 0.07, 0.06, 0.06)$.

a) Find a binary Huffman code for X. Compute its expected length.

b) Find a ternary Huffman code for X. Compute its expected length.

Problem 2 — *Bad Codes*

Which of these codes cannot be Huffman codes for any probability assignment?

a) $\{0, 10, 11\}$.

b) $\{00, 01, 10, 110\}$.

c) $\{01, 10\}$.

Problem 3 — *Optimal Codeword Lengths*

Although the codeword lengths of an optimal prefix-free code are complicated functions of the message probabilities $\{p_1, p_2, \ldots, p_m\}$, it can be said that less probable symbols are encoded into longer codewords. Suppose that the message probabilities are given in decreasing order $p_1 > p_2 \geq \cdots \geq p_m$.

a) Prove that for any binary Huffman code, if the most probable message symbol has probability $p_1 > \frac{2}{5}$, then that symbol must be assigned a codeword of length 1.

b) Prove that for any binary Huffman code, if the most probable message symbol has probability $p_1 < \frac{1}{3}$, then that symbol must be assigned a codeword of length at least 2.

Problem 4 ***AEP***

Let X_1, X_2, ... be drawn IID according to the probability mass function $p(x)$, $x \in \{1, 2, \dots, m\}$. Thus, $p(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} p(x_i)$. We know that

$$-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) \to H(X)$$

in probability. Let $q(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} q(x_i)$, where q is another probability mass function on $\{1, 2, \dots, m\}$.

a) Evaluate $\lim_{n\to\infty} -\frac{1}{n} \log q(X_1, X_2, \dots, X_n)$ and express it with entropies and relative entropies.

b) Evaluate $\lim_{n\to\infty} \frac{1}{n} \log \frac{p(X_1, X_2, \dots, X_n)}{q(X_1, X_2, \dots, X_n)}$.

Problem 5 ***Proof of Theorem 3.3.1 in Cover & Thomas: High-Probability Sets and the Typical Set***

Let $X_1, \dots, X_n$ be IID $\sim P_X$. For some fixed $\epsilon > 0$, let $\mathcal{A}_\epsilon^{(n)}$ denote the set of weakly typical sequences, and for some fixed $\delta > 0$, let $\mathcal{B}_\delta^{(n)}$ be an arbitrary set of length-n sequences such that

$$\Pr\big(\mathcal{B}_\delta^{(n)}\big) \geq 1 - \delta.$$

a) Given any two sets $\mathcal{A}$, $\mathcal{B}$ such that $\Pr(\mathcal{A}) \geq 1 - \epsilon_1$ and $\Pr(\mathcal{B}) \geq 1 - \epsilon_2$, show that

$$\Pr(\mathcal{A} \cap \mathcal{B}) \geq 1 - \epsilon_1 - \epsilon_2.$$

For n sufficiently large, show that this implies

$$\Pr\big(\mathcal{A}_\epsilon^{(n)} \cap \mathcal{B}_\delta^{(n)}\big) \geq 1 - \epsilon - \delta.$$

b) For n sufficiently large, justify the following steps:

$$\begin{aligned}
1 - \epsilon - \delta &\overset{\text{(i)}}{\leq} \Pr\big(\mathcal{A}_\epsilon^{(n)} \cap \mathcal{B}_\delta^{(n)}\big) \\
&\overset{\text{(ii)}}{=} \sum_{\mathbf{x} \in \mathcal{A}_\epsilon^{(n)} \cap \mathcal{B}_\delta^{(n)}} P_{\mathbf{X}}(\mathbf{x}) \\
&\overset{\text{(iii)}}{\leq} \sum_{\mathbf{x} \in \mathcal{A}_\epsilon^{(n)} \cap \mathcal{B}_\delta^{(n)}} 2^{-n(H(X)-\epsilon)} \\
&\overset{\text{(iv)}}{=} \big|\mathcal{A}_\epsilon^{(n)} \cap \mathcal{B}_\delta^{(n)}\big| \cdot 2^{-n(H(X)-\epsilon)} \\
&\overset{\text{(v)}}{\leq} \big|\mathcal{B}_\delta^{(n)}\big| \cdot 2^{-n(H(X)-\epsilon)}.
\end{aligned}$$

c) Show that for any $\delta \in (0, 1)$, any $\delta' > 0$, and n sufficiently large,

$$\frac{1}{n} \log\big|\mathcal{B}_\delta^{(n)}\big| \geq H(X) - \delta'.$$