E2 201 (Aug–Dec 2014)

Homework Assignment 4

Discussion: Friday, Sept. 26 Quiz: TBA

This assignment consists of two pages.

1. [Problem 5.36, Cover & Thomas, 2nd ed.]
   (a) Can $(1, 2, 2)$ be the codeword lengths of a binary Huffman code? What about $(2, 2, 3, 3)$?
   (b) What codeword lengths $(\ell_1, \ell_2, \ldots, \ell_m)$ can arise from binary Huffman codes?

2. [Problem 5.17(b), Cover & Thomas, 2nd ed.]
   Find an optimal binary prefix-free code for the probability mass function (p_k) defined by

$$p_k = (2/3)\,(1/3)^{k-1}, \;\; k = 1, 2, 3, \ldots.$$

3. Problem 5.20, Cover & Thomas, 2nd ed.

4. Problem 5.25, Cover & Thomas, 2nd ed.

5. *Rényi entropy*.
   Let X be a discrete random variable with pmf $p(x), x \in \mathcal{X}$. For any real number $\alpha > 0$, $\alpha \neq 1$, the Rényi entropy of order α is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \sum_{x \in \mathcal{X}} p(x)^\alpha.$$

   As usual, $H(X)$ denotes the Shannon entropy $-\sum_x p(x) \log_2 p(x)$.

   (a) Show that $H_\alpha(X) \geq H(X)$ for all $\alpha < 1$, and as $\alpha \to 1$, $H_\alpha(X)$ converges to $H(X)$.

   Let $a > 1$ be a real number. For a source code $\mathcal{C} = \{c(x) : x \in \mathcal{X}\}$ with codeword lengths $\ell(x), x \in \mathcal{X}$, define the *$a$-exponential mean* of $\mathcal{C}$ to be

$$L_a(\mathcal{C}) = \log_a \mathbb{E}[a^{\ell(X)}] = \log_a \left(\sum_{x \in \mathcal{X}} p(x) a^{\ell(x)} \right)$$

   (b) Recall that $L(\mathcal{C})$ is the expected codeword length: $L(\mathcal{C}) = \mathbb{E}[\ell(X)]$. Show that $L_a(\mathcal{C}) \geq L(\mathcal{C})$ for all $a > 1$, and as $a \to 1$, $L_a(\mathcal{C})$ converges to $L(\mathcal{C})$.

   (c) Define $\alpha = \frac{1}{1+\log_2 a}$, and note that $\alpha \to 1$ as $a \to 1$. For $x \in \mathcal{X}$, set

$$\ell(x) = \left\lceil -\log_2 \frac{p(x)^\alpha}{\sum_{x \in \mathcal{X}} p(x)^\alpha} \right\rceil$$

   Show that there exists a prefix code $\mathcal{C}$ with codeword lengths $\ell(x), x \in \mathcal{X}$, and furthermore,

$$H_\alpha(X) \leq L_a(\mathcal{C}) < H_\alpha(X) + 1.$$

   (d) From the result of part (b), deduce that for any uniquely decodable code $\mathcal{C}$,

$$L_a(\mathcal{C}) \geq H(X).$$

   [*Remark*: It can in fact be shown (by means of Hölder's inequality) that for any uniquely decodable code $\mathcal{C}$, $L_a(\mathcal{C}) \geq H_\alpha(X)$ with $\alpha = \frac{1}{1+\log_2 a}$. Thus, for an optimal code $\mathcal{C}^*$ (which minimizes $L_a(\mathcal{C})$ among all uniquely decodable codes $\mathcal{C}$), we have $H_\alpha(X) \leq L_a(\mathcal{C}^*) < H_\alpha(X) + 1$.]

The definition of Rényi entropy is due to Alfred Rényi (1961). The results of parts (b)–(d) above are due to Lorne Campbell (1966).

6. (a) Let Y be a continuous random variable with cumulative distribution function (cdf) $F(y)$. Show that $W = F(Y)$ is uniformly distributed over $[0, 1]$.

   (b) Let W be a continuous random variable uniformly distributed over the interval $[0, 1]$. Consider the binary representation of W as $0.X_1X_2X_3\ldots$. Show that $X_1, X_2, X_3, \ldots$ form a sequence of i.i.d. Bernoulli$(1/2)$ random variables.

7. Let $X_1, X_2, X_3, \ldots$ be any sequence of binary random variables specified by the joint distributions $p(x^n)$, $n = 1, 2, \ldots$. Let F denote the cumulative distribution function of $Y = \sum_{i=1}^{\infty} X_i 2^{-i}$ (i.e., Y is the real number with binary representation $0.X_1X_2X_3\ldots$). Show that if $y \in [0, 1]$ has binary representation $0.x_1x_2x_3\ldots$, then

$$F(y) = \sum_{k:x_k=1} p(x^{k-1}0).$$

8. Consider a binary stationary Markov source $\mathbf{X}$ with transition probability matrix

$$P = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

   Length-3 sequences (X_1, X_2, X_3) generated by this source are to be encoded into binary codewords using a lossless source code.

   (a) Compute the entropy rate $\mathcal{H}(\mathbf{X})$, and the entropy $H(X_1, X_2, X_3)$.

   (b) Determine an optimal binary prefix-free code for encoding length-3 sequences (X_1, X_2, X_3) generated by the source. What is the expected codeword length?

   (c) If the length-3 sequences generated by the source were to be encoded using arithmetic coding, what would be the expected codeword length?

   (d) Determine the codewords assigned by arithmetic coding to the source sequences $(0, 1, 1)$ and $(1, 0, 0)$.