ECE 745 - Advanced Communication Theory, Spring 2007
Homework #1

Due: March 12, 2007

1. A Huffman code finds the optimal codeword to assign to a given block of N source symbols.

   (a) Suppose that the probability of the most probable block is greater than 0.4; show that this implies that the codeword generated by a Huffman code for this block has length 1.

   (b) Suppose that the probability of the most probable block is less than $\frac{1}{3}$; show that this implies that the codeword generated by a Huffman code for this block has length ≥ 2.

   (c) Show that $\{01, 100, 101, 1110, 1111, 0011, 0001\}$ cannot be a Huffman code, regardless of the source distribution and the choice of N.

2. (a) Describe a pair of random variables X and Y such that there exists y_1 and y_2 such that

$$H(X|Y = y_1) \leq H(X) \leq H(X|Y = y_2)$$

   (b) Describe a pair of random variables X and Y for which $I(X;Y) = H(X) < H(Y)$.

   (c) Let $Z = X + Y$ for real-valued discrete random variables X and Y. Show that $H(Z) \leq H(X) + H(Y)$. Give an example where $H(Z) = H(X) + H(Y)$ and another where $H(Z) < H(X) + H(Y)$.

   (d) Show that if X and Z are conditionally independent given Y, then

$$H(X|Z) \geq H(X|Y)$$

3. For a given independent and identically distributed (IID) discrete source, the best lossless source code for $N = 2$ has average length 4 and the best lossless source code for $N = 3$ has average length 4.8.

   (a) What can be deduced about the entropy of the source?

   (b) What can be deduced about the size of the source alphabet?

4. Define the *variational distance* between probability mass functions $p(x)$ and $q(x)$, each with alphabet $\mathcal{X}$ as

$$V(p,q) = \sum_{\mathcal{X}} |p(x) - q(x)|$$

   (a) Show that $V(p,q)$ has all of the properties of a metric; that is, show (i) $V(p,q) \geq 0$ with equality if and only if $p = q$, (ii) $V(p,q) = V(q,p)$, and (iii) $V(p,q) \leq V(p,r) + V(r,q)$ for any three probability mass functions p, q, r, each with alphabet $\mathcal{X}$.

   (b) For $|\mathcal{X}| < \infty$, show that for any probability mass function $p(x)$ and any $\epsilon > 0$, there exists a $\delta > 0$ so small that $D(p||q) \leq \epsilon$ whenever $V(p,q) \leq \delta$.

5. A discrete-valued source outputs an independent and identically distributed (IID) sequence of random variables (X_i), each drawn from the alphabet $\mathcal{X} = \{a_1, a_2, \ldots, a_Q\}$. I have not been able to find the probability mass function $p_{X_i}(x)$; however, I have been able to show that for any $\epsilon > 0$, there are sets $B_1, B_2, B_3, \ldots$ that have the following three properties:

   (i) B_N is a subset of $\mathcal{X}^N$.

   (ii) $P((X_1, X_2, \ldots, X_N)^T \in B_N) \to 1$ as $N \to \infty$.

   (iii) $|B_N| = 4^{3\epsilon N + 5N}$.

   What can be said about the entropy of the source?

6. The following is a formal restatement of our theorem for almost lossless fixed length source codes:

   Let (X_i) be an IID sequence of discrete random variables with common distribution $p_X(x)$ and first-order entropy $H(X)$.

   (1) Then, for any $R > H(X)$, any $\delta > 0$, there exists a fixed length source code with rate less than R and probability of decoding error $P_e < \delta$.

   (2) Conversely, for any $R < H(X)$ and any $\epsilon > 0$, there is a (large) integer N_0 such that for all $N \geq N_0$, a fixed length source code that operates on N symbols at a time has probability of decoding error bounded as:

$$P_e > 1 - \epsilon - 2^{N(R - H(X) + \epsilon)}$$

   By applying the Typical Sequences Theorem, show:

   (a) The positive statement (statement (1)).

   (b) The converse (statement (2)).

   (c) Show that for $R < H(X)$, $\lim_{N \to \infty} P_e = 1$.

7. Suppose that we are trying to "guess" X from an observation of Y with the decision rule $\hat{X}(Y)$. Let $P_e = P(\hat{X}(Y) \neq X)$.

   (a) Show that any optimum decision rule can never have P_e greater than $\frac{|\mathcal{X}|-1}{|\mathcal{X}|}$.

   (b) Show that Fano's function $\mathcal{F}_{|\mathcal{X}|}(p)$ is maximized at $p = \frac{|\mathcal{X}|-1}{|\mathcal{X}|}$.

   (c) Suppose $I(X;Y) = H(X)$. Show that an optimal decision rule has $P_e = 0$.