EE 8510 Thursday, Jan. 27, 2005
Advanced Topics in Communications Prof. N. Jindal

Homework Set # 2
Due: Thursday, Feb. 3, 2005

1. *Entropy of functions of a random variable.* Let X be a discrete random variable. Show that the entropy of a function of X is less than or equal to the entropy of X by justifying the following steps:

$$\begin{aligned} H(X, g(X)) &\stackrel{(a)}{=} H(X) + H(g(X)|X) \\ &\stackrel{(b)}{=} H(X) \\ H(X, g(X)) &\stackrel{(c)}{=} H(g(X)) + H(X|g(X)) \\ &\stackrel{(d)}{\geq} H(g(X)) \end{aligned}$$

Thus $H((g(X)) \leq H(X)$.
(Cover & Thomas 2.5)

2. *An AEP-like limit.* Let $X_1, X_2, \ldots$ be i.i.d. drawn according to probability mass function $p(x)$. Find

$$\lim_{n\to\infty} [p(X_1, X_2, \ldots, X_n)]^{1/n}$$

(Cover & Thomas 3.2)

3. (Csiszar, Korner 78) Let X^n and Y^n be two random vectors with arbitrary joint probability distribution. Show that

$$\sum_{i=1}^{n} I(X_{i+1}^n; Y_i|Y^{i-1}) = \sum_{i=1}^{n} I(Y^{i-1}; X_i|X_{i+1}^n) \quad \forall 1 \leq i \leq n$$

where $X_{n+1}, Y_0 = \emptyset$. We use the following notation: $X^j = (X_1, \ldots, X_j)$ and $X_k^j = (X_k, \ldots, X_j)$. Note that this expression may be useful later in proving some converses.

4. In this problem we will show that capacity is the fundamental limit on communication even if the performance criterion is bit error rate instead of block error rate. In lecture we were interested in systems that had block error rate P_B converging to zero as block length (n) went to infinity. However, in many practical systems we are interested in the bit error rate P_b.

   Consider a system where the message W is an equiprobable k-bit sequence, denoted $U_1, \ldots, U_k$, and the output of the encoder is an n-bit sequence denoted $X_1, \ldots, X_n$. Clearly the rate of our code is $R = k/n$. The output of the channel is $Y_1, \ldots, Y_n$. The decoder outputs an estimate of the message bits denoted $\hat{U}_1, \ldots, \hat{U}_k$.

   The expressions for block and bit error rate are:

$$P_B = \Pr\{(\hat{U}_1, \ldots, \hat{U}_k) \neq U_1, \ldots, U_k\} \qquad P_b = \frac{1}{k}\sum_{i=1}^{k} \Pr\{\hat{U}_i \neq U_i\}$$

   It is easy to see that $P_b \leq P_B$ since each block error rate can cause at most k bit errors (i.e. in the worst case each bit is wrong).

   (a) Use Fano's inequality to prove

$$H(U_i|\hat{U}_i) \leq h(P_{e,i})$$

   where $h(\cdot)$ is the binary entropy function and $P_{e,i} = \Pr\{\hat{U}_i \neq U_i\}$.

   (b) Use the concavity of the entropy function to show

$$\frac{1}{k}\sum_{i=1}^{k} h(P_{e,i}) \leq h\left(\frac{1}{k}\sum_{i=1}^{k} P_{e,i}\right) = h(P_b)$$

   (c) Justify steps (a)-(g) in the following proof:

$$\begin{aligned} k = H(U^k) &\overset{(a)}{=} I(U^k; \hat{U}^k) + H(U^k|\hat{U}^k) \\ &\overset{(b)}{\leq} I(X^n; Y^n) + H(U^k|\hat{U}^k) \end{aligned}$$

$$\begin{aligned}
&\overset{(c)}{\leq} \sum_{i=1}^{n} I(X_i;Y_i) + H(U^k|\hat{U}^k) \\
&\overset{(d)}{\leq} nC + H(U^k|\hat{U}^k) \\
&\overset{(e)}{\leq} nC + \sum_{i=1}^{k} H(U_i|\hat{U}_i) \\
&\overset{(f)}{\leq} nC + \sum_{i=1}^{k} h(P_{e,i}) \\
&\overset{(g)}{\leq} nC + kh(P_b)
\end{aligned}$$

Dividing both sides by k and rearranging gives:

$$h(P_b) \geq 1 - \frac{C}{R} \quad \rightarrow \quad P_b \geq h^{-1}\left(1 - \frac{C}{R}\right).$$

Since the $h(p) > 0$ for $0 < p < 1$, the bit error rate is bounded away from zero for all blocklengths n if $R > C$.

For example, if you are using a code with $R = 1.1C$, you are guaranteed that $P_b \geq h^{-1}\left(1 - \frac{1}{1.1}\right) = h^{-1}(0.091) = .0115$, i.e. a BER of larger than 1%. (J. Massey notes)

5. *An additive noise channel.* Find the channel capacity of the following discrete memoryless channel: $Y = X + Z$, where $\Pr(Z = 0) = \Pr(Z = a) = \frac{1}{2}$. The alphabet for x is $\mathcal{X} = \{0, 1\}$. Assume that Z is independent of X. Observe that the channel capacity depends on the value of a. (Cover & Thomas 8.3)

6. *Cascade of binary symmetric channels.* Show that a cascade of n identical binary symmetric channels,

$$X_0 \rightarrow \text{BSC \#1} \rightarrow X_1 \rightarrow \cdots \rightarrow X_{n-1} \rightarrow \text{BSC \#n} \rightarrow X_n$$

each with raw error probability p, is equivalent to a single BSC with error probability $\frac{1}{2}(1 - (1 - 2p)^n)$ and hence $\lim_{n\to\infty} I(X_0; X_n) = 0$ if $p \neq 0, 1$. No encoding or decoding takes place at the intermediate terminals $X_1, \ldots, X_{n-1}$. Thus the capacity of the cascade tends to zero. (Cover & Thomas 8.8)

7. *The Z channel.* The Z channel has binary input and output alphabets and transition probabilities $p(y|x)$ given by $p(y = 0|x = 0) = 1$ and $p(y = 0|x = 1) = p(y = 1|x = 1) = \frac{1}{2}$. Find the capacity of the Z channel and the maximizing input probability distribution. (Cover & Thomas 8.9)