EE606: Final Exam

(Total 100 points, Allowed time: 2 hours)

Given: December 17, 2004

1. *(15 pts)* Random variable X has as its range the set of non-negative integers. Find the probability assignment $P_X(n)$, $n = 0, 1, 2, \ldots$, that maximizes $H(X)$ subject to the constraint that $E(X) = a$. Evaluate the resulting $H(X)$.

   Solution

   The Lagrangian for the problem is given by

   $$L(p_0, p_1, \ldots, \lambda, \mu) = -\sum_{n=0}^{\infty} p_n \log p_n + \lambda\Big(\sum_{n=0}^{\infty} np_n - a\Big) + \mu\Big(\sum_{n=0}^{\infty} p_n - 1\Big).$$

   Assume that log's are to base e. We have

   $$\frac{\partial L}{\partial p_n} = -\log p_n - p_n \frac{1}{p_n} + n\lambda + \mu$$

   Setting $\frac{\partial L}{\partial p_n} = 0$ for $n = 0, 1, \ldots$ gives

   $$\boxed{p_n = e^{n\lambda + \mu - 1}}$$

   Constraint $\sum p_n = 1$ gives $e^{\mu-1} = (1 - e^{\lambda})$, and constraint $\sum np_n = a$ gives $e^{\mu-1}e^{\lambda} = a(1 - e^{\lambda})$. Therefore,

   $$p_n = \frac{a^n}{(1+a)^{n+1}},$$

   giving $H = (1 + a)\log(1 + a) - a \log a$.

2. *(Total 25 pts)* Consider a discrete memoryless channel with input X and output Y, both taking values from the alphabet {0, 1, 2, 3, 4, 5, 6, 7}. The probability transition matrix $p(y|x)$ is given by

   $$p(y|x) = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 0 & 1/4 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 0 & 0 & 1/4 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 0 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 1/4 & 0 & 0 & 1/4 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1/4 & 0 & 1/4 & 1/4 & 1/4 \end{bmatrix}$$

   (a) *(5 pts)* Find the mutual information $I(X;Y)$ when X is uniformly distributed.

   (b) *(5 pts)* Find the channel capacity and a capacity achieving distribution.

   (c) *(5 pts)* Find the mutual information $I(X;Y)$ when X is distributed as follows:

   $$p(x) = \begin{cases} 1/2, & \text{if } x \in \{0, 7\}, \\ 0, & \text{otherwise.} \end{cases}$$

(d) *(10 pts)* Demonstrate a capacity achieving code.

Solution

(a) Computing by definition gives $I(X,Y) = 1$.

(b) Note that the channel is symmetric. Therefore $C = \log 8 - H(1/4, 1/4, 1/4, 1/4) = 1$, and is achieved by the uniform distribution on the input.

(c) Computing by definition gives $I(X,Y) = 1$. This distribution is also capacity achieving.

(d) Recall the noisy typewriter channel and see (b). The following rate 1 code is error-free:
Encoder:

$$0 \to 0$$
$$0 \to 7$$

Decoder:

$$\{0,1,2,4\} \to 0$$
$$\{3,5,6,7\} \to 1$$

Therefore this code is capacity achieving.

3. *(Total 40 pts)* Consider a communications system with two sensors S_X and S_Y. S_X observes random variable X, and S_Y, random variable Y. Values of both X and Y are required at the same destination. The joint probability distribution P_{XY} is given by the following matrix:

X \ Y	000	001	010	011	100	101	110	111
000	p	p	p	0	p	0	0	0
001	p	p	0	p	0	p	0	0
010	p	0	p	p	0	0	p	0
011	0	p	p	p	0	0	0	p
100	p	0	0	0	p	p	p	0
101	0	p	0	0	p	p	0	p
110	0	0	p	0	p	0	p	p
111	0	0	0	p	0	p	p	p

$p = 1/32$

(a) *(10 pts)* Consider the case when only sensor S_X is in service. What is the minimum number of bits that S_X has to send to the destination to report the exact value of X it measured.

(b) *(10 pts)* Consider the case when both sensors are in service and can cooperate. If they use a single link to the destination, what is the minimum number of bits that must be sent through the link to report the exact value of (X,Y).

(c) *(10 pts)* Consider the case when both sensors are in service but cannot cooperate. What is the minimum number of bits that the destination needs to receive from each of the sensors to be able to recover the exact value of (X,Y) they measured.

(d) *(10 pts)* Demonstrate a coding scheme that could be used to achieve the theoretical bounds you found in (c).

Solution

(a) There are 8 equally probable values that X can take. Therefore we need $H(X) = 3$ bits to describe it. Note that $H(Y) = 3$ as well.

(b) There are 32 equally probable values that (X, Y) can take. Therefore we need $H(X, Y) = 5$ bits to describe it.

(c) By the Slepian-Wolf theorem, we know that

$$\begin{aligned} R_X &\geq H(X|Y) \\ R_Y &\geq H(Y|X) \\ R_X + R_Y &\geq H(X, Y) \end{aligned}$$

Since $H(X|Y) = H(X, Y) - H(Y) = 2$, and by symmetry, $H(Y|X) = 2$ as well, each sensor has to send at least 2 bits and the sum of their rates has to be at least 5.

(d) Consider the following coloring of X:

X \ Y	000	001	010	011	100	101	110	111
000	p	p	p	0	p	0	0	0
001	p	p	0	p	0	p	0	0
010	p	0	p	p	0	0	p	0
011	0	p	p	p	0	0	0	p
100	p	0	0	0	p	p	p	0
101	0	p	0	0	p	p	0	p
110	0	0	p	0	p	0	p	p
111	0	0	0	p	0	p	p	p

$p = 1/32$

Look at any two rows of the same color. Note that an observation Y can occur in only one of those rows. Let sensor S_Y send 3 bits identifying observation Y unambiguously. Then it is enough for sensor S_X to send the two bits identifying the subset it belongs to.

subset:	{000, 111}	{001, 110}	{010, 101}	{100, 011}
label:	00	01	10	11

The knowledge of Y at the receiver will resolve the ambiguity within the subset.

4. *(10 pts)* Consider the Gaussian channel

$$Y = X + Z,$$

where $Z \sim \mathcal{N}(0, \sigma^2)$, Z is independent of the input X. Instead of the usual input power constraint, we now have an output power constraint P, that is, $E(Y^2) \leq P$. Find the capacity of this channel.

Solution

We have derived that

$$I(X;Y) = h(Y) - h(Z) = h(Y) - \frac{1}{2}\log 2\pi e\sigma^2, \tag{1}$$

and shown that, if $E(Y^2) \leq P$, then

$$h(Y) \leq \frac{1}{2}\log 2\pi eP. \tag{2}$$

Combining (1) and (2) gives

$$I(X;Y) \leq \boxed{\frac{1}{2}\log\frac{P}{\sigma^2}}.$$

Since the equality can be achieved with $X \sim \mathcal{N}(0, P-\sigma^2)$, we have found the capacity.

5. *(10 pts)* Suppose we have a DMC with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and transition probabilities $\{p(y|x) : y \in \mathcal{Y}, x \in \mathcal{X}\}$. We define a single-letter distortion measure between a source letter $x \in \mathcal{X}$ and a reproduction letter $x' \in \mathcal{X}$ as

$$d(x,x') = -\log \sum_{y\in\mathcal{Y}} \sqrt{p(y|x)p(y|x')}, \quad x, x' \in \mathcal{X},$$

and distortion between sequences as

$$d(\boldsymbol{x},\boldsymbol{x'}) = -\frac{1}{n}\log \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sqrt{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x'})}, \quad \boldsymbol{x}, \boldsymbol{x'} \in \mathcal{X}^n.$$

Show that the distortion between sequences is the average of the per letter distortion between elements of the sequences, namely that the following holds:

$$d(\boldsymbol{x},\boldsymbol{x'}) = \frac{1}{n}\sum_{i=1}^{n} d(x_i, x_i').$$

Solution

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} d(x_i,x_i') &= -\frac{1}{n}\sum_{i=1}^{n}\log\sum_{y\in\mathcal{Y}}\sqrt{p(y|x)p(y|x')} \\
&= -\frac{1}{n}\log\prod_{i=1}^{n}\Big[\sum_{y\in\mathcal{Y}}\sqrt{p(y|x_i)p(b|x_i')}\Big] \\
&= -\frac{1}{n}\log\sum_{\boldsymbol{y}\in\mathcal{Y}^n}\sqrt{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x'})} \\
&= d(\boldsymbol{x},\boldsymbol{x'})
\end{aligned}$$