ECE 776 - Information theory (Spring 2010)
Final

Q1 (1 point). Consider a Markov chain $X_1 - X_2 - X_3$ with $|\mathcal{X}_1| = 3$, $|\mathcal{X}_2| = 2$ $|\mathcal{X}_3| = 4$ prove that $I(X_1; X_3) \leq 1$.

Sol.: We have

$$I(X_1; X_3) \leq I(X_1; X_2) \leq H(X_2) \leq \log_2 |\mathcal{X}_2| = 1,$$

where the first inequality follows from the data processing inequality.

Q2 (1 point). Consider an i.i.d. process X^n with pdf $f(x) = c \cdot 2^{-x^4}$, where c is an appropriate constant. Prove that

$$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} x_i^4 \leq h(X) + \log_2 c + \epsilon\right] \rightarrow 1 \text{ as } n \rightarrow \infty$$

for any $\epsilon > 0$ (Hint: What is the set of typical sequences?)

Sol.: The set of typical sequences is

$$A_\epsilon^{(n)} = \left\{x^n \in \mathbb{R}^n\colon\ h(X) - \epsilon \leq -\frac{1}{n}\log_2 f(x^n) \leq h(X) + \epsilon\right\}$$

but $f(x^n) = c^n \prod_{i=1}^n 2^{-x_i^4}$, so $-\frac{1}{n}\log_2 f(x^n) = -\log_2 c + \frac{1}{n}\sum_{i=1}^n x_i^4$, so that we can write

$$A_\epsilon^{(n)} = \left\{x^n \in \mathbb{R}^n\colon\ \log_2 c + h(X) - \epsilon \leq \frac{1}{n}\sum_{i=1}^{n} x_i^4 \leq \log_2 c + h(X) + \epsilon\right\},$$

which answers the question.

Q3 (1 point). Consider an additive white Gaussian noise channel $Y = X + Z$ with $Z \sim \mathcal{N}(0, N)$ subject only to an output power constraint of P (no constraint is imposed on the input!). Evaluate the corresponding capacity

$$C = \max_{\substack{f(x):\\ E[Y^2] \leq P}} I(X; Y)$$

Sol.: We have

$$\begin{aligned}
I(X;Y) &= h(Y) - h(Y|X) = h(Y) - \frac{1}{2}\log_2(2\pi e N) \\
&\leq \frac{1}{2}\log_2(2\pi e P) - \frac{1}{2}\log_2(2\pi e N) \\
&= \frac{1}{2}\log_2\left(\frac{P}{N}\right),
\end{aligned}$$

where the inequality follows from the fact that the entropy $h(Y)$ is maximized by a Gaussian distribution under power constraint $E[Y^2] \leq P$.

Q4 (1 point). Consider the i.i.d. Gaussian vector process $[X_1\ X_2]$ with zero-mean, independent entries and powers $E[X_1^2] = 1$ and $E[X_2^2] = 3$. Focusing on the MSE distortion, find the rate-distortion function using reverse waterfilling.

Sol.: Using reverse waterfilling, we find that

- If $D \geq 4$, $R(D) = 0$;
- If $2 \leq D \leq 4$, $R(D) = \frac{1}{2}\log_2 \frac{3}{D-1}$;
- If $D \leq 2$, $R(D) = \frac{1}{2}\log_2 \frac{1}{D/2} + \frac{1}{2}\log_2 \frac{3}{D/2}$.

P1 (2 points) Consider the channel

$$p(y|x) = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

where rows correspond to different $x \in \{0, 1, 2\}$ and columns to different $y \in \{0, 1, 2\}$. The number of channel uses is n.
a. Argue that symbol $x = 2$ should not be used.
b. Using only the input set $\{0, 1\}$ find the optimal input distribution and the capacity.
c. Assume that we want to operate the channel with a (suboptimal) code (of length n) that is randomly generated with $p_X(0) = 3/4$ and $p_X(1) = 1/4$. What is the rate achievable with this code?
d. To achieve the rate R obtained at the previous point, consider the following transmission scheme. An i.i.d. binary code B^{2n} of length $2n$ and rate $2R$ is constructed i.i.d. with $p_B(0) = 1/2$ and $p_B(1) = 1/2$. Every two bits of B^{2n} are mapped into one input bit of the transmitted sequence X^n. Find one such mapping so that X^n has the desired distribution ($p_X(0) = 3/4$ and $p_X(1) = 1/4$).

Sol.: a. Transmitting $x = 2$ may result in any output with the same probability, so that the decoder has no means to find out whether $x = 2$ was transmitted.
b. The channel is (weakly) symmetric. In fact, it is an erasure channel with erasure probability 1/3. Therefore, the optimal input distribution is $p_X(0) = 1/2$. The capacity is

$$C = 1 - 1/3 = 2/3.$$

c. The rate achievable by this code is

$$\begin{aligned} R &= I(X;Y) \\ &= H(Y) - H(1/3) \\ &= H(1/2, 1/3, 1/6) - 0.92 \\ &= 1.46 - 0.92 = 0.54, \end{aligned}$$

where we have calculated $p_Y(0) = 3/4 \cdot 2/3 = 1/2$, $p_Y(1) = 1/3 \cdot 3/4 + 1/3 \cdot 1/4 = 1/3$, $p_Y(2) = 1/6$.
d. A mapping that satisfies the requirement is

$$\begin{aligned} 00 &\rightarrow 0 \\ 01 &\rightarrow 0 \\ 10 &\rightarrow 0 \\ 11 &\rightarrow 1. \end{aligned}$$

P2 (2 points) Consider a horse race with four horses with winning probabilities $p = \{1/8, 1/8, 1/4, 1/2\}$. The odds fair with respect to the uniform distribution $r = \{1/4, 1/4, 1/4, 1/4\}$. You start with \$1.
a. Consider the betting strategy $b = \{1/4, 1/4, 1/4, 1/4\}$. What is the (long-term) doubling rate of this strategy? Approximately, how much money you would have after 100 races?
b. What is the optimal betting strategy b_1, b_2, b_3, b_4 that maximizes the (long-term) doubling rate? Find the corresponding doubling rate. Approximately, how much money you would have after 100 races?

Sol.: a. The doubling rate is

$$\begin{aligned} W &= D(p||r) - D(p||b) \\ &= \sum_k p_k \log_2 \left(\frac{b_k}{r_k}\right) = 0, \end{aligned}$$

therefore the amount of money after 100 races is about

$$S_n \simeq \$1 \cdot 2^{nW} = \$1.$$

b. The optimal betting strategy is $b = p$. This leads to the doubling rate

$$\begin{aligned} W^* &= D(p||r) \\ &= 1/4 \end{aligned}$$

so that

$$\begin{aligned} S_n &\simeq \$1 \cdot 2^{\frac{100}{4}} \\ &= \$33554432. \end{aligned}$$

P3 (2 points) Consider the additive Gaussian noise channel

$$Y_i = X_i + Z_i,$$

where we have power constraint on X^n equal to $P = 3$. The noise Z^n has independent Gaussian samples but with possibly non-zero and time-varying mean.
a. Assume that $Z_i \sim \mathcal{N}\left(\frac{1}{i^2}, 1\right)$ and find the capacity.
b. Assume that $Z_i \sim \mathcal{N}(Q_i, 1)$ with Q_i being a random variable with $Q_i \sim \mathcal{N}(0, 2)$. Find the capacity.

c. For the scenario at point a, consider the transmission of an i.i.d. Gaussian process V^n with $V_i \sim \mathcal{N}(0,1)$. Can we construct a source-channel coding scheme that enables transmission at MSE equal to 0.75?
d. Repeat the point above with MSE equal to 0.2.

Sol.: a. By evaluating the mutual information, one obtains $C = 1/2 \log_2(1+3) = 1$.
b. We can write

$$Z_i = Q_i + U_i,$$

where $U_i \sim \mathcal{N}(0,1)$ and independent of Q_i. Therefore, $Z_i \sim \mathcal{N}(0,3)$ and $C = 1/2 \log_2(1 + 3/3) = 1/2$.
c. We know that this is possible if and only if

$$R(D) = \frac{1}{2}\log_2 \frac{1}{D} \leq C = 1.$$

Therefore, $D = 0.75$ is feasible since $R(0.75) = 0.2$.
d. This is instead not possible since $R(0.2) = 1.16$.