

ECE 520.447
Introduction to Information Theory and Coding

Midterm Examination #2

1:00 — 2:00 PM, April 17, 2000.

Name: _____

Read these instructions before starting the examination.

- (i) This is a closed-book examination. Use of the textbook, notes, *etc.*, is not permitted. Some useful formulae appear at the end of the examination booklet.
- (ii) Use of electronic calculators is permitted for numeric calculations only.
- (iii) Show all your work clearly and concisely. Points may be deducted for illegible or unclear answers.
- (iv) Provide answers in the space provided. Use the unprinted side of the pages in the examination booklet if necessary.
- (v) There are three mandatory questions for a total of 50 points. You must answer **all** parts of these questions. There is an optional 10-point *bonus question*. Points earned on the bonus question will be added to your total. Students enrolled in the Ph.D. program are *strongly encouraged* to attempt the bonus question.

Best of luck!

Question No 1	/10 Points
Question No 2	/20 Points
Question No 3	/20 Points
Bonus Question	/10 Points

TOTAL	/50 Points
-------	------------

Question No 1: Consider a 3 horse race with win probabilities

$$\mathbf{p} = (p_1, p_2, p_3) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right).$$

Assume that a bookmaker offers *o*-for-1 odds on this race with $\mathbf{o} = (o_1, o_2, o_3) = (4, 4, 2)$.

(1a) What is the log-optimal betting strategy $\mathbf{b}^* = (b_1^*, b_2^*, b_3^*)$ for this race, and what is the doubling rate of \mathbf{b}^* (2 points)

(1b) Find the *set* of *all* bets \mathbf{b} for which the compounded wealth in repeated plays will grow to infinity. You may wish to note that $\sum_{i=1}^3 \frac{1}{o_i} = 1$ and use the notation $\mathbf{r} = (r_1, r_2, r_3) = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right)$. (5 points)

(1c) What odds $\mathbf{o}' = (o'_1, o'_2, o'_3)$ *minimize* the chances of a gambler if the bookmaker is required by law to have $\sum_{i=1}^3 \frac{1}{o'_i} \leq 1$? (3 points)

Question No 2: Let $\mathcal{X} = \mathcal{Y} = \{0, 1, 2, 3\}$ be the input and output alphabet(s) of a discrete memoryless channel with a transition probability matrix

$$W = \begin{bmatrix} 1 - \epsilon & \epsilon & 0 & 0 \\ \epsilon & 1 - \epsilon & 0 & 0 \\ 0 & 0 & 1 - \delta & \delta \\ 0 & 0 & \delta & 1 - \delta \end{bmatrix}.$$

Such a channel is sometimes called a *sum channel*, because it may be thought of as the “sum” or union of two parallel subchannels

$$W_1 = \begin{bmatrix} 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \end{bmatrix} \quad \text{and} \quad W_2 = \begin{bmatrix} 1 - \delta & \delta \\ \delta & 1 - \delta \end{bmatrix}.$$

with alphabets $\mathcal{X}_1 = \mathcal{Y}_1 = \{0, 1\}$ and $\mathcal{X}_2 = \mathcal{Y}_2 = \{2, 3\}$ respectively. This problem addresses the capacity of a sum channel.

(2a) “Draw” the channel transition diagram for W . (2 points)

(2b) Compute the capacity of the channel W for $\epsilon = \delta = \frac{1}{2}$. (4 points)

(2c) Let $p(x)$ be a probability mass function on \mathcal{X} . Let

$$p(0) + p(1) = \alpha \quad \text{and} \quad p(2) + p(3) = 1 - \alpha.$$

Show that the mutual information between the input X and the output Y of the channel W may be written as (4 points)

$$I(X; Y) = H(\alpha) + \alpha I(X; Y | X \in \mathcal{X}_1) + (1 - \alpha) I(X; Y | X \in \mathcal{X}_2)$$

(2d) To compute the capacity of the channel W , we must maximize the mutual information in 2(c) over all choices of $p(x)$. Argue why this may be done in two steps by first *separately* choosing conditional probabilities $p_1(x|X \in \mathcal{X}_1)$ and $p_2(x|X \in \mathcal{X}_2)$ to maximize $I(X; Y|X \in \mathcal{X}_1)$ and $I(X; Y|X \in \mathcal{X}_2)$ respectively, and then choosing α to maximize $I(X; Y)$. (2 points)

(2e) Let C_1 and C_2 denote the capacity of the subchannels W_1 and W_2 respectively. Argue why $\max_{p(x)} I(X; Y) = \max_{\alpha} H(\alpha) + \alpha C_1 + (1 - \alpha)C_2$. (2 points)

(2f) Show that the capacity C of the sum channel W is given by

$$C = \log(2^{C_1} + 2^{C_2}),$$

where C_1 and C_2 are the capacities of the subchannels W_1 and W_2 . (6 points)

Question No 3: (Converse to the Channel Capacity Theorem for the discrete time Gaussian Channel) Assume that a sequence of $(2^{nR}, n)$ channel codes for a discrete-time memoryless Gaussian channel

$$Y_i = X_i + Z_i, \quad i = 1, \dots, n,$$

with power constraint P has the property that $P_e \rightarrow 0$. In this problem, we will reconstruct arguments for why its rate R must satisfy

$$R \leq C = \frac{1}{2} \log \left(1 + \frac{P}{N} \right).$$

To begin with, let the codewords for this “good” code be arranged in the form of a matrix with $M = 2^{nR}$ rows as

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_n(1) \\ x_1(2) & x_2(2) & \cdots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(M) & x_2(M) & \cdots & x_n(M) \end{bmatrix},$$

and note that the power constraint implies that

$$\frac{1}{n} \sum_{i=1}^n x_i^2(m) \leq P, \quad \text{for each } m = 1, \dots, M.$$

Let us denote the “column” sums of squares by

$$P_i = \frac{1}{M} \sum_{m=1}^M x_i^2(m), \quad \text{for each } i = 1, \dots, n.$$

Finally, let W be a random variable which has a uniform distribution over the message set $\{1, 2, \dots, M\}$.

The proof of the converse proceeds according to the following inequalities:

$$\begin{aligned} nR = H(W) &= I(W; Y^n) + H(W|Y^n) \\ &\leq I(W; Y^n) + 1 + nRP_e \end{aligned} \tag{1}$$

$$\leq I(X^n; Y^n) + 1 + nRP_e \tag{2}$$

$$\begin{aligned} &= h(Y^n) - h(Y^n|X^n) + 1 + nRP_e \\ &= h(Y^n) - h(Z^n) + 1 + nRP_e \end{aligned} \tag{3}$$

$$\leq \sum_{i=1}^n h(Y_i) - h(Z^n) + 1 + nRP_e \tag{4}$$

$$= \sum_{i=1}^n (h(Y_i) - h(Z_i)) + 1 + nRP_e \tag{5}$$

$$\leq \sum_{i=1}^n \left(h(Y_i) - \frac{1}{2} \log(2\pi eN) \right) + 1 + nRP_e \tag{6}$$

$$\leq \sum_{i=1}^n \left(\frac{1}{2} \log(2\pi e(P_i + N)) - \frac{1}{2} \log(2\pi eN) \right) + 1 + nRP_e \tag{7}$$

$$= \sum_{i=1}^n \left(\frac{1}{2} \log \left(1 + \frac{P_i}{N} \right) + \frac{1}{n} + RP_e \right)$$

$$\begin{aligned}
R &\leq \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{2} \log \left(1 + \frac{P_i}{N} \right) \right) + \frac{1}{n} + RP_e \\
&\leq \left(\frac{1}{2} \log \frac{1}{n} \sum_{i=1}^n \left(1 + \frac{P_i}{N} \right) \right) + \frac{1}{n} + RP_e \tag{8}
\end{aligned}$$

$$= \frac{1}{2} \log \left(1 + \frac{P}{N} \right) + \frac{1}{n} + RP_e \tag{9}$$

From this, it is clear that if $P_e \rightarrow 0$ then $R \leq C$.

(3a) Provide justification for all the numbered (in)equalities in the derivation shown above.
(10 points)

(1)

(2)

(3)

(4)

(5)

(6)

(7)

(8)

(9)

- (3b) Let X be a \mathbb{R} -valued randomvariable with density $f(x)$. From first principles, derive the differential entropy of a random variable $Y = aX + c$ in terms of the differential entropy of X for some constants c and $a \neq 0$. (10 points)

Bonus Question: Suppose you wish to design a (n, k) channel code with minimum distance $2t+1$ (or greater), where k is the number of message bits and $n-k$ is the number of additional or redundant parity-check bits. You would of course want $n-k$ to be as small as possible.

(4a) Show that for *any* binary (n, k) linear code with minimum distance $2t+1$,

$$n - k \geq \log_2 \left[1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{t} \right].$$

This inequality can alternately be viewed as an upper-bound on the error correcting capability t of an (n, k) code. (7 points)

(4b) Show that Hamming codes achieve this bound with equality. (3 points)

This bound is known as the *Hamming bound* and was derived by F. J. MacWilliams in 1963.

Extra Work Space