ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 35 | Information Theory and Coding

Final exam | Jan. 18, 2020

4 problems, 100 points, each part is worth 5 points
180 minutes
2 sheets (4 pages) of notes allowed.

Good Luck!

PLEASE WRITE YOUR NAME ON EACH SHEET OF YOUR ANSWERS.

PLEASE WRITE THE SOLUTION OF EACH PROBLEM ON A SEPARATE SHEET.

Problem 1. (20 points) Suppose $X_1, ..., X_n$ are discrete random variables. For a subset $\mathcal{S}$ of $\{1, \dots, n\}$ define $X_{\mathcal{S}} = \{X_i : i \in \mathcal{S}\}$. For $k = 1, \dots, n$, define

$$H_k = \frac{1}{\binom{n}{k}} \sum_{\mathcal{S}:|\mathcal{S}|=k} H(X_{\mathcal{S}}),$$

as the average entropy of all possible k-tuples of random variables X_i's, e.g., $H_1 = \frac{H(X_1)+\cdots+H(X_n)}{n}$, $H_n = H(X_1, \dots, X_n)$. We define $H_0 = 0$.

a) Show that for any three random variables A, B, C,

$$H(AB) + H(BC) \geq H(B) + H(ABC).$$

b) Show that for any subsets $\mathcal{S}$ and $\mathcal{T}$ of $\{1, \dots, n\}$,

$$H(X_{\mathcal{S}}) + H(X_{\mathcal{T}}) \geq H(X_{\mathcal{S}\cup\mathcal{T}}) + H(X_{\mathcal{S}\cap\mathcal{T}}).$$

[Hint: use (a) with a suitable choice of A, B, C.]

c) Show that $H_k - H_{k-1} \geq H_{k+1} - H_k$.

[Hint: The left-hand side is the average of $H(X_{i_{k+1}}|X_{i_2}, \dots, X_{i_k})$ over all permutations $(i_1, \dots, i_n)$ of $(1, \dots, n)$. Write the right-hand side as a similar average and compare each term.]

d) Show that

$$\frac{H_{k+1}}{k+1} \leq \frac{H_k}{k}.$$

[Hint: $H_k = \sum_{i=1}^{k}(H_i - H_{i-1})$.]

PROBLEM 2. (25 points) Let W be a random variable with $P(W = 0) = P(W = 1) = P(W = 2) = 1/3$. The random process $Z_1, Z_2, \ldots$ is defined conditioned on W as follows:

- if $W = 0$ or $W = 1$ then $Z_i = W$ for all i,
- if $W = 2$, then Z_i's are i.i.d. with $\Pr(Z_i = 1|W = 2) = \Pr(Z_i = 0|W = 2) = 1/2$.

Observe that $Z_1, Z_2, \ldots$ is a stationary process (since for any k the statistics of $Z_k, Z_{k+1}, \ldots$ are the same as the process $Z_1, Z_2, \ldots$).

a) Find the entropy rate $H_Z := \lim_{n\to\infty} H(Z_1^n)/n$. [Hint: Consider $H(Z_1^n, W)$ and $H(Z_1^n|W)$.]

Suppose we have a binary-input binary-output communication channel, whose input $x_1, x_2, \ldots$, and output $Y_1, Y_2, \ldots$ are related via $Y_i = x_i + Z_i \mod 2$.

b) Define $C_n = \max_{p_{X^n}} I(X^n; Y^n)/n$. Show that

$$C_n = 1 - H(Z_1^n)/n.$$

What is the p_{X^n} that achieves this equality?

c) What is $C = \lim_{n\to\infty} C_n$?

Suppose that we attempt to send one bit of information over this channel by designing a block code of blocklength n.

d) Show that the error probability of any code is at least $1/6$.

e) What is the capacity of this channel?

PROBLEM 3. (30 points) A binary code is said to be a *constant-weight code* if the Hamming weights of all codewords are the same. From any binary code $\mathcal{C}$ of blocklength n, we can create a constant-weight code $\mathcal{C}_k$ for $k \in \{0, \dots, n\}$ by only taking the codewords with Hamming weight k, i.e. $\mathcal{C}_k = \{c \in \mathcal{C} : w_H(c) = k\}$.

Let the message set be $\mathcal{U}$. Given any encoder and decoder pair (enc, dec) for $\mathcal{C}$ on channel W, we will denote the maximum error probability as

$$p_e(enc, dec) := \max_{u \in \mathcal{U}} W^n(dec(Y^n) \neq u | X^n = enc(u)).$$

a) Show that the capacity of a channel W can be achieved by constant-weight codes. [Hint: For any code $\mathcal{C}$ of rate R and error probability p_e, show that there is a $\mathcal{C}_k$ with rate $R' \geq R - \frac{\log_2(n+1)}{n}$ and $p'_e \leq p_e$.]

For any code $\mathcal{C}$ with encoder enc on channel W, we can define an erasure-only decoder

$$dec_{eo}(y^n) = \begin{cases} u, & W^n(y^n|enc(u)) > 0 \text{ and } \forall_{u' \neq u} W^n(y^n|enc(u')) = 0 \\ ? & \text{else.} \end{cases}$$

This decoder only decides if $T(y^n) := \big|\{u : W(y^n|enc(u)) > 0\}\big|$, the number of compatible codewords is equal to 1, i.e., if it is sure of making a correct decision. For any $\mathcal{C}$ with encoder enc, we define the erasure probability $p_{eo}(enc, dec_{eo}) := \max_{u \in \mathcal{U}} \Pr(dec_{eo}(Y^n) = ?|enc(u))$.

The erasures-only capacity of W, $C_{eo}(W)$, is the supremum of rates R such that for any $\epsilon > 0$ there is a code $\mathcal{C}$ with rate R and $p_{eo} < \epsilon$.

b) What is C_{eo} of a Binary Symmetric Channel?

Fix p in $[0, 1)$ and suppose for the rest of the problem that $W = \text{BEC}(p)$ (*not* $\text{BSC}(p)$). For the rest of the problem, we consider a constant-weight code $\mathcal{C}_k$ with (enc_k, dec_k).

c) For any $c \in \mathcal{C}_k$ and y^n containing j erasures, show that $W^n(y^n|c)$ is equal to either zero or $p^j(1-p)^{n-j}$.

We now assume that each message is chosen with equal probability, i.e., $U = u$ with equal probability for all $u \in \mathcal{U}$.

d) Define $B = \{y^n : T(y^n) > 1\}$. Show that $\Pr(U = u|Y^n = y^n) \leq 1/2$ for all $y^n \in B$ and $u \in \mathcal{U}$.

e) Set $\hat{U} = dec_k(Y^n)$. Show that $\Pr(\hat{U} \neq U) \geq \frac{1}{2}\Pr(Y^n \in B)$. [Hint: For $y^n \in B$, what does part (d) imply about $\Pr(\hat{U} = U|Y^n = y^n)$.]

f) Show that for any $R < C(W)$ and $\epsilon > 0$, there is a constant-weight code $\mathcal{C}_k$ with rate at least R such that $p_{eo} < \epsilon$. What is the relation between $C_{eo}(W)$ and $C(W)$?

PROBLEM 4. (25 points) A binary code is said to be a k *constant-weight* code if the Hamming weight of all codewords are equal to k, $k > 0$. We want to determine the maximum number of codewords of a k constant-weight code with blocklength n and minimum distance d.

Assume that we have a k constant-weight code $\mathcal{C}_k$ with $M \geq 2$ codewords and encoder enc, i.e. for all $i \in \{1, \dots, M\}$ we have $w_H(enc(i)) = k$.

We also define $x_{i,j}$ as the value of $enc(i)$ at the j-th coordinate, e.g. if $enc(2) = 00101$ then $x_{2,5} = 1$ and $x_{2,4} = 0$. Also define $w_j = \sum_{i=1}^{M} x_{i,j}$, i.e. the number of codewords which have 1 at the j-th coordinate. Note that the addition and multiplication of $x_{i,j}$ and w_j is the standard addition and multiplication on real numbers, (instead of the binary addition and multiplication).

a) Can d be an odd number? Does there exist a constant-weight code which is also a linear code? Justify your answer.

b) Show that for any $a \neq b$,
$$\sum_{j=1}^{n} x_{a,j} x_{b,j} \leq k - \frac{d}{2}.$$

c) Show that
$$\frac{k^2 M^2}{n} \leq \sum_{j=1}^{n} w_j^2.$$
[Hint : $\sum_{j=1}^{n} w_j = kM$.]

d) Show that
$$\frac{k^2 M}{n} - k \leq (M-1)\left(k - \frac{d}{2}\right).$$

e) Define $M^*(n, d, k)$ as the maximum number of codewords that a k constant-weight code with blocklength n and minimal distance d can have. Show that $M^*(9, 6, 4) = 3$.