ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 19 Information Theory and Coding
Homework 8 Nov. 12, 2019

PROBLEM 1. Consider a memoryless channel with transition probability matrix $P_{Y|X}(y|x)$, with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. For a distribution Q over $\mathcal{X}$, let $I(Q)$ denote the mutual information between the input and the output of the channel when the input distribution is Q. Show that for any two distributions Q and Q' over $\mathcal{X}$,

(a)
$$I(Q') \leq \sum_{x\in\mathcal{X}} Q'(x) \sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) \log\left(\frac{P_{Y|X}(y|x)}{\sum_{x'\in\mathcal{X}} P_{Y|X}(y|x')Q(x')}\right)$$

(b)
$$C \leq \max_x \sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) \log\left(\frac{P_{Y|X}(y|x)}{\sum_{x'\in\mathcal{X}} P_{Y|X}(y|x')Q(x')}\right)$$

where C is the capacity of the channel. Notice that this upper bound to the capacity is independent of the maximizing distribution.

PROBLEM 2.

(a) Show that $I(U;V) \geq I(U;V|T)$ if T, U, V form a Markov chain, i.e., conditional on U, the random variables T and V are independent.

Fix a conditional probability distribution $p(y|x)$, and suppose $p_1(x)$ and $p_2(x)$ are two probability distributions on $\mathcal{X}$.

For $k \in \{1,2\}$, let I_k denote the mutual information between X and Y when the distribution of X is $p_k(\cdot)$.

For $0 \leq \lambda \leq 1$, let W be a random variable, taking values in $\{1,2\}$, with

$$\Pr(W=1) = \lambda, \quad \Pr(W=2) = 1-\lambda.$$

Define

$$p_{W,X,Y}(w,x,y) = \begin{cases} \lambda p_1(x)p(y|x) & \text{if } w=1 \\ (1-\lambda)p_2(x)p(y|x) & \text{if } w=2. \end{cases}$$

(b) Express $I(X;Y|W)$ in terms of I_1, I_2 and λ.

(c) Express $p(x)$ in terms of $p_1(x)$, $p_2(x)$ and λ.

(d) Using (a), (b) and (c) show that, for every fixed conditional distribution $p_{Y|X}$, the mutual information $I(X;Y)$ is a concave $\cap$ function of p_X.

PROBLEM 3. Consider a random source $\mathcal{S}$ of information, and let W be a random variable which represents the first L symbols $U_1, \ldots, U_L$ of this source, i.e., $W = U_1^L$. We want to transmit the value of W using a memoryless stationary channel as follows:

- At time $t=1$, we send $X_1 = f_1(W)$ through the channel.

- At time $t = i + 1 \leq n$, we send $X_{i+1} = f_i(W, Y^i)$ through the channel. $Y_1, \ldots, Y_i$ are the output of the channel at times $t = 1, \ldots, i$ respectively,

$f_1, \ldots, f_n$ are n mappings that constitute the encoder. Clearly, this is a communication system with feedback as we are using the value of Y^i in the computation of X_{i+1}.

In the previous problem, we gave an example which satisfies $I(X^n; Y^n) > nC$ and $I(W; Y^n) \leq nC$. Show that the inequality $I(W; Y^n) \leq nC$ always holds by justifying each of the following equalities and inequalities:

$$\begin{aligned} I(W; Y^n) &\overset{(a)}{=} \sum_{i=1}^{n} I(W; Y_i | Y^{i-1}) \overset{(b)}{\leq} \sum_{i=1}^{n} I(W, Y^{i-1}; Y_i) \overset{(c)}{\leq} \sum_{i=1}^{n} I(W, X_i, X^{i-1}, Y^{i-1}; Y_i) \\ &\overset{(d)}{=} \sum_{i=1}^{n} I(X_i, X^{i-1}, Y^{i-1}; Y_i) \overset{(e)}{=} \sum_{i=1}^{n} I(X_i; Y_i) \overset{(f)}{\leq} nC. \end{aligned}$$

Since $I(W; Y^n)$ represents the amount of information that is shared with the receiver, the inequality $I(W; Y^n) \leq nC$ shows that feedback does not increase the capacity.

Problem 4. Suppose Z is uniformly distributed on $[-1, 1]$, and X is a random variable, independent of Z, constrained to take values in $[-1, 1]$. What distribution for X maximizes the entropy of $X + Z$? What distribution of X maximizes the entropy of XZ?

Problem 5. Random variables X and Y are correlated Gaussian variables:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, K = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} \right).$$

Find $I(X; Y)$.

Problem 6. Suppose X and Y are independent geometric random variables. That is, $p_X(k) = (1-p)^{k-1}p$ and $p_Y(k) = (1-q)^{k-1}q, \quad \forall k \in \{1, 2, \ldots\}$.

(a) Find $H(X, Y)$.

(b) Find $H(2X + Y, X - 2Y)$

Now consider two independent exponential random variables X and Y. That is, $p_X(t) = \lambda_X e^{-\lambda_X t}$ and $p_Y(t) = \lambda_Y e^{-\lambda_Y t}, \quad \forall t \in [0, \infty)$.

(c) Find $h(X, Y)$.

(d) Find $h(2X + Y, X - 2Y)$