Georgia Tech

School of Mathematics
Math 4280

Information Theory
Final Exam
April 29 2016, 11 :30am-2 :20pm

First Name : ____

Name : ____

DO NOT WRITE IN THE TABLE BELOW

1.a		3.a		5.a	
1.b		3.b		5.b	
1.c		3.c		5.c	
1.d		3.d		5.d	
2.a		3.e		5.e	
2.b		4.a		5.f	
2.c		4.b		5.g	
		4.c			

WARNING :

Read carefully, read the comments in italic, take your time, do not panic and double check what you write.

Write CLEARLY your answer where it is asked to.

The final exam will last 2 hours and 50 minutes.

1. Let X, Y be two random variables on a finite probability space, which will be denoted by $\Omega_X \times \Omega_Y$.

    (a) Give the definition of the mutual information $I(X;Y)$.

    (b) Express the mutual information in terms of the entropies $H(X,Y), H(X), H(Y)$ and also in terms of $H(X), H(X|Y)$.

    (c) Prove that $I(Y;X) = 0$ if and only if X and Y are independent.

    (d) Interpret the definition of the mutual information in terms of the Kullback-Leibler distance.

(Use this page to finish your proof if necessary)

2. Let $\xi = (X_n)_{n=1}^{\infty}$ be a stochastic process made of variables on a finite probability spaces Ω.

   (a) What does it mean for this process to be stationary ?

   (b) Let $h_n = H(X_n|X_1, \cdots, X_{n-1})$. Show that if ξ is stationary, the sequence $(h_n)_{n=1}^{\infty}$ is positive and non-increasing.

   (c) Deduce from the previous question that, if ξ is stationary, the following limit exist (called the entropy rate)

$$h(\xi) = \lim_{n\to\infty} \frac{1}{n} H(X_1, \cdots, X_n) .$$

(Use this page to finish your proof if necessary)

3. Let $X^n = (X_1, X_2, \cdots, X_n)$ be *i.i.d.*'s drawn from a finite probability space (Ω, q) and with common entropy $H(X)$ and common average $\mu = \mathbb{E}(X)$. Let A_ϵ^n and B_ϵ^n be the sets of sequences $x^n = (x_1, \cdots, x_n) \in \Omega^n$ defined by

$$A_\epsilon^n = \left\{ x^n \in \Omega^n \,;\, \left| -\frac{1}{n} \log p(x^n) - H(X) \right| \leq \epsilon \right\},$$

$$B_\epsilon^n = \left\{ x^n \in \Omega^n \,;\, \left| \frac{1}{n} \sum_{k=1}^{n} x_k - \mu \right| \leq \epsilon \right\},$$

where $p(x^n) = \text{Prob}\{X^n = x^n\}$.

(a) Give the expression of the joint probability distribution of $X^n = (X_1, X_2, \cdots, X_n)$ in terms of the individual probability q; compute its logarithm.

(b) Does $\text{Prob}\{X^n \in A_\epsilon^n\} \to 1$ as $n \to \infty$? Give an argument to justify your answer.

(c) Does $\text{Prob}\{X^n \in A_\epsilon^n \cap B_\epsilon^n\} \to 1$ as $n \to \infty$? Give an argument to justify your answer.

(d) Show that the number of elements of $A_\epsilon^n \cap B_\epsilon^n$ is less than or equal to $2^{n(H(X)+\epsilon)}$ for all n.

(e) Show that the number of elements of $A_\epsilon^n \cap B_\epsilon^n$ is larger than or equal to $(1/2)2^{n(H(X)-\epsilon)}$ for n large enough.

(Use this page to finish your proof if necessary)

4. Let $\widehat{X}$ be a random variable drawn from the set $\Omega = \{0, 1, 2, \cdots, 9, 10\}$ having 11 points. Let also Z taking on one of the three values $\{1, 2, 3\}$ with equal probability. Let the channel taking $\widehat{X}$ to $\widehat{Y} = \widehat{X} + Z \,(\text{mod} 11)$ be considered here.

   (a) Give the precise general definition of the channel capacity for a channel with input X and output Y.

   (b) Compute the channel capacity for the channel described above.

   (c) Give the optimal probability distribution of $\widehat{X}$ reaching this capacity.

(Use this page to finish your proof if necessary)

5. **Reminder :** *Let X be a random variable on the finite set Ω with distribution $p(x) = \text{Prob}\{X = x\}$. Let $\widehat{X}$ be another random variable on Ω. We set $q(\widehat{x}|x) = \text{Prob}\{\widehat{X} = \widehat{x}|X = x)$ and $r(\widehat{x}) = \text{Prob}\{\widehat{X} = \widehat{x}\}$. A distortion function $d(x, \widehat{x}) \geq 0$ is introduced to evaluate how well $\widehat{X}$ represents X.*

   During the semester, it was proved that if q minimizes the mutual information $I(X; \widehat{X})$ with the distortion constraint $D = \mathbb{E}(d(X, \widehat{X}))$, then it is given by the following formula

$$q(\widehat{x}|x) = \frac{r(\widehat{x})\, e^{-\lambda d(x,\widehat{x})}}{Z(x)}\,, \qquad Z(x) = \sum_{\widehat{y}\in\Omega} r(\widehat{y})\, e^{-\lambda d(x,\widehat{y})}\,, \tag{1}$$

   where λ is a Lagrange multiplier to be fixed by the constraint. In addition the definition of r implies $\sum_x p(x)q(\widehat{x}|x) = r(\widehat{x})$ leading to the identity

$$1 = \sum_{x\in\Omega} \frac{p(x)e^{-\lambda d(x,\widehat{x})}}{Z(x)}\,. \tag{2}$$

   The rate-distortion function $R(D)$ is given by $I(X; \widehat{X})$ computed with this distribution when the Lagrange multiplier method applies. It might be zero otherwise. It has been proved that $R(D)$ is both non increasing and convex as a function of D.

   Problem : Let X be a uniformly distributed over $\Omega = \{1, 2, \cdots, m\}$, where $m > 2$. The Hamming distortion is defined by

$$d(x, \widehat{x}) = \begin{cases} 0 & \text{if } x = \widehat{x} \\ 1 & \text{if } x \neq \widehat{x} \end{cases}$$

   The goal of this problem is to compute the rate-distortion function $R(D)$ in this case.

   (a) Show that $0 \leq D < 1$.

   (b) Show that $Z(x) = \{e^{-\lambda}/m\}\,\{r(x)(e^{\lambda} - 1) + m - 1\}$.

   (c) Use eq. (2) to show that $\widehat{X}$ is uniformly distributed
   (Hint : set $S = \sum_x 1/Z(x)$ to simplify notations and conclude that $r(\widehat{x})$ is independent of $\widehat{x}$.)

   (d) Using eq. (1), give the expression of $q(\widehat{x}|x)$ in terms of m and e^{λ}.

   (e) Compute D and express e^{λ} in terms of m and d.

   (f) Compute the mutual information $I(X; \widehat{X})$ with the previous solution, in order to get $R(D)$.

   (g) Show that whenever $D = (m - 1)/m$, both $R(D)$ and its first derivative vanish. Conclude that the previous formula is valid in the interval $0 \leq D \leq (m - 1)/m$ but not beyond. What is $R(D)$ if $D > (m - 1)/m$?

(Use this page to finish your proof if necessary)

(Use this page to finish your proof if necessary)

(Use this page to finish your proof if necessary)