MTHE - MATH 474/874 - Information Theory
Fall 2019

A Sample of Past Final Problems

1. Decide whether each of the following statements is *true* or *false*. Prove the validity of those that are true and give counterexamples or arguments based on known facts to disprove those that are false.

   (a) Let X and Y be two jointly distributed discrete random variables with alphabet $\mathcal{X}$ and $\mathcal{Y}$, respectively. Define the random variable Z by $Z = g(X)$, where $g(\cdot)$ is an arbitrary function with domain $\mathcal{X}$. Then $H(Y|X) \leq H(Y|Z)$.

   (b) Every Huffman code for a discrete memoryless source (DMS) has a corresponding suffix code with the same average code rate.

   (c) There exists a binary erasure channel with capacity $C = \sqrt{\pi}/2$ bits/channel use.

   (d) There exists a uniformly distributed random variable X with differential entropy $h(X) = -3$ bits.

   (e) Random coding is a method of proving the existence of codes with certain desirable properties.

   (f) The discrete entropy function $H(X)$ of a random variable X with finite alphabet $\mathcal{X}$ and pmf $p(\cdot)$ is concave in $p(\cdot)$.

   (g) The differential entropy $h(X)$ is invariant under invertible transformations; i.e., if g is invertible then $h(X) = h(g(X))$.

   (h) The rate of an n-th order Huffman code for a DMS $\{X_i\}_1^\infty$ is arbitrarily close to the source entropy $H(X)$ iff $n \to \infty$.

   (i) Among all probability density functions with support the interval $[a, b]$, the uniform density function has the largest differential entropy.

   (j) The capacity of a DMC with finite input and output alphabets is always bounded from above.

(k) Suppose that random variables X, Y and Z are jointly Gaussian, each with mean 0 and variance 1. Assume that $X \to Y \to Z$ and that $E[XY] = \rho$, where $0 < \rho < 1$. Then

$$I(X;Z) > \frac{1}{2}\log_2\left[\frac{1}{1-\rho^2}\right].$$

(l) Let random variable X be uniformly distributed over the interval $[0,1]$, and define random variable Y by $Y = 2X + 2$. Then $h(Y) < 0$.

(m) For a stationary source $\{X_n\}_{n=1}^{\infty}$, its normalized joint entropy $\frac{1}{n}H(X_1, \cdots, X_n)$ is nonincreasing in n.

2. Answer the following questions.

(a) Let random variables X and Y have Gaussian joint distribution, with $E[X] = 0$, $E[Y] = 0$, $E[X^2] = \sigma_1^2$, $E[Y^2] = \sigma_2^2$, and $E[XY] = \sigma_1\sigma_2\rho$, where $0 \le \rho < 1$. Find $I(X;Y)$ in terms of ρ. For what value of ρ are X and Y independent ?

(b) A discrete-time memoryless Gaussian channel has equal input and noise powers. Find the capacity of this channel in bits/channel use and explain its operational significance.

3. The lognormal density of a positive random variable X is given by

$$f_X(x) = \frac{1}{\sigma x\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \quad , x > 0, \quad -\infty < \mu < +\infty, \quad , \sigma > 0.$$

Compute the differential entropy $h(X)$ in nats.

4. (a) State the Data Processing Theorem and prove it.

(b) The output of a discrete memoryless channel K_1 is connected to the input of another discrete memoryless channel K_2. Show that the capacity of the cascade combination can never exceed the capacity of K_i, $i = 1, 2$.

5. Let X, Y and Z denote three discrete random variables. Show that

$$H(X|Y) + H(Y|Z) \ge H(X|Z).$$

6. Consider a discrete-time binary stationary Markov source $\{U_L\}_{L=1}^{\infty}$ with transition probability

$$Pr\{U_{L+1} = j | U_L = i\} = \begin{cases} 0.05 & \text{if } j = 1 \text{ and } i = 0, \\ 0.50 & \text{if } j = 0 \text{ and } i = 1. \end{cases}$$

   (a) Compute the source entropy rate $H(\mathcal{U})$.

   (b) Is it possible to *reliably* transmit this Markov source across a discrete memoryless channel with capacity $C = 0.5$ bits/channel use ? Explain qualitatively. (Assume that the source and channel signaling rates are equal.)

7. Let X be a binary random variable with alphabet $\mathcal{X} = \{0, 1\}$. Let Z denote another random variable that is independent of X and taking values in $\mathcal{Z} = \{0, 1, 2, 3\}$ such that $\Pr\{Z = 0\} = \Pr\{Z = 1\} = \Pr\{Z = 2\} = \epsilon$, where $0 < \epsilon < 1$. Consider a discrete memoryless channel with input X, noise Z, and output Y described by the equation:

$$Y = 3X + (-1)^X Z,$$

   where X and Z are as defined above.

   (a) Determine the channel transition probability matrix $Q \triangleq [p(y|x)]$, and draw the channel transition diagram.

   (b) Compute the capacity C of this channel in terms of ϵ. What is the maximizing input distribution that achieves capacity ?

   (c) For what value of ϵ is the noise entropy $H(Z)$ maximized ? What is the value of C for this choice of ϵ ? Comment on the result.

8. (a) Find the entropy rate in bits per source symbol of the binary stationary Markov chain with transition matrix

$$P = \begin{bmatrix} 2/3 & 1/3 \\ 1/5 & 4/5 \end{bmatrix}.$$

   (b) Design a second-order binary Shannon-Fano code for the source. Is it optimal ?

(c) For what values of p, can you transmit reliably the source over a channel with the following matrix

$$Q = \begin{bmatrix} \frac{1-p}{2} & \frac{1-p}{2} & \frac{p}{2} & \frac{p}{2} \\ \frac{p}{2} & \frac{p}{2} & \frac{1-p}{2} & \frac{1-p}{2} \end{bmatrix}.$$

9. Let $\{X_i\}_{i=1}^{\infty}$ be a stationary source with alphabet $\mathcal{X}$. Define a new source $\{Y_i\}_{i=1}^{\infty}$ with alphabet $\mathcal{Y}$ by $Y_i = f(X_i)$ where $f(\cdot)$ is a given function. Compare the entropy rates of the two sources – i.e., find the relationship between $H(\mathcal{X})$ and $H(\mathcal{Y})$.

10. *Inequalities.* Which of the following inequalities are $\geq$, $=$, $\leq$? Label each with $\geq$, $=$, or $\leq$ and justify your answer.

    (a) $I(g(X); Y)$ versus $I(X; Y)$ where $g(\cdot)$ is a given function.

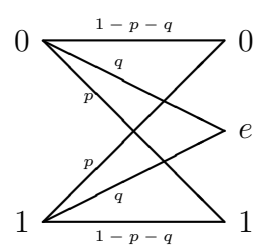
    (b) $H(X, Y)/(H(X) + H(Y))$ versus 1.

    (c) $H(X_2|X_1)$ versus $H(X_3|X_2, X_1)$ for a stationary source $\{X_n\}_{n=1}^{\infty}$.

    (d) $H(X_3|X_2)$ versus $\lim_{n\to\infty} H(X_n|X_{n-1}, \cdots, X_1)$ for a stationary Markov source $\{X_n\}_{n=1}^{\infty}$.

    (e) $h(4X + 2)$ versus $h(X)$ for a real-valued random variable X.

11. *Binary-Input Channel with Erasures and Errors.* Consider the following two-input, three-output discrete memoryless channel



where $p \in (0, 1)$, $q \in (0, 1)$ and $p + q \in (0, 1)$.

(a) Find the capacity C of this channel in terms of p and q. What is the maximizing input distribution that achieves capacity ?

(b) If $p = q = 1/3$, find the value of C. Comment on the result.

12. (a) *Fading channel.* Consider an additive noise fading channel described by:

$$Y = XV + Z,$$

where Y is the channel output, X is the channel input, V is a random variable representing fading, and Z is additive noise. Assume that X, Z and V are all independent of each other. Argue that knowledge of the fading factor V improves capacity by showing that

$$I(X;Y|V) \geq I(X;Y).$$

(b) Incidentally, note that conditioning does not always increase mutual information. Give an example of jointly distribution random variables U, R and S such that

$$I(U;R|S) < I(U;R).$$

13. Prove by using Jensen's inequality that the divergence $D(p\|q) \geq 0$, where $p(\cdot)$ and $q(\cdot)$ are two probability mass functions defined on the same alphabet.

14. *The Sum Channel.*

(a) Let $(\mathcal{X}_1, \mathcal{Y}_1, p_1(y|x))$ be a discrete memoryless channel (DMC) with finite input alphabet $\mathcal{X}_1$, finite output alphabet $\mathcal{Y}_1$, transition distribution $p_1(y|x)$ and capacity C_1. Similarly, let $(\mathcal{X}_2, \mathcal{Y}_2, p_2(y|x))$ be another DMC with capacity C_2. Assume that $\mathcal{X}_1 \cap \mathcal{X}_2 = \emptyset$ and that $\mathcal{Y}_1 \cap \mathcal{Y}_2 = \emptyset$.

Now let $(\mathcal{X}, \mathcal{Y}, p(y|x))$ be the *sum* of these two channels where $\mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2$, $\mathcal{Y} = \mathcal{Y}_1 \cup \mathcal{Y}_2$ and

$$p(y|x) = \begin{cases} p_1(y|x) & \text{if } x \in \mathcal{X}_1, y \in \mathcal{Y}_1 \\ p_2(y|x) & \text{if } x \in \mathcal{X}_2, y \in \mathcal{Y}_2 \\ 0 & \text{otherwise.} \end{cases}$$

Show that the capacity of the sum channel is given by

$$C_{sum} = \log_2 \left[2^{C_1} + 2^{C_2}\right] \qquad \text{bits/channel use.}$$

(b) Compute C_{sum} above if the first channel is a binary symmetric channel with crossover probability 0.11, and the second channel is a binary erasure channel with erasure probability 0.5.

15. Given a fixed positive integer $n > 1$, consider an n-ary valued random variable X with alphabet $\mathcal{X} = \{1, 2, \cdots, n\}$ and distribution described by the probabilities $p_i \triangleq Pr\{X = i\}$, where $p_i > 0$ for each $i = 1, \cdots, n$. Given $\alpha > 0$ and $\alpha \neq 1$, define the Rényi entropy of X as

$$H_\alpha(X) \triangleq \frac{1}{1-\alpha} \log_2 \left[\sum_{i=1}^{n} p_i^\alpha \right].$$

(a) Show that

$$\sum_{i=1}^{n} p_i^r > 1 \quad \text{if} \quad r < 1,$$

and that

$$\sum_{i=1}^{n} p_i^r < 1 \quad \text{if} \quad r > 1.$$

[*Hint*: Show that the function $f(r) = \sum_{i=1}^{n} p_i^r$ is decreasing in r, where $r > 0$.]

(b) Show that

$$0 \leq H_\alpha(X) \leq \log_2 n.$$

[*Hint*: Use (a) for the lower bound, and use Jensen's inequality (with the convex function $f(y) = y^{\frac{1}{1-\alpha}}$, for $y > 0$) for the upper bound.]

(c) Find the limit of $H_\alpha(X)$ as $\alpha \to 1$. Comment on the result.

16. Answer the following:

(a) State the channel coding theorem.

(b) Discuss briefly two main tools used in proving the forward part of the channel coding theorem.

(c) State the main inequality used in proving the converse part.

17. State the lossless joint source-channel coding theorem for a memoryless source-channel pair, and explain why it can be used to justify the separate design of the source and channel coding operations in a communication system.

18. Let X be a discrete random variable with alphabet $\mathcal{X}$ and distribution $p(x)$. Let $f : \mathcal{X} \to \mathbb{R}$ be a real-valued function, and let α be an arbitrary real number.

(a) Show that

$$H(X) \leq \alpha E[f(X))] + \log_2 \left[\sum_{x \in \mathcal{X}} 2^{-\alpha f(x)} \right],$$

with equality iff $p(x) = \frac{1}{A} 2^{-\alpha f(x)}$, where $A \triangleq \sum_{x \in \mathcal{X}} 2^{-\alpha f(x)}$.

(b) Show that for a positive integer-valued random variable N, the following hold:

$$H(N) \leq \log_2(E[N]) + \log_2 e.$$

Hint: First use part (a) with $f(N) = N$ and $\alpha = \log_2 \frac{E[N]}{E[N]-1}$; then use the fact that $\sum_{n=1}^{\infty} a^n = \frac{a}{1-a}$ if $|a| < 1$ to obtain that

$$H(N) \leq E[N] \log_2 \frac{E[N]}{E[N]-1} + \log_2(E[N]-1) = (E[N]-1) \log_2 \frac{E[N]}{E[N]-1} + \log_2 E[N],$$

and apply the fundamental inequality.
