EE 231A
Information Theory
Instructor: Rick Wesel

Handout #14, Problem Set 5
Tuesday April 28, 2020
Due: Tuesday, May 5, 2020

94 pts
Reading: Chapters 8 & 9 of *Elements of Information Theory*

Lecture 9: Fano's Inequality and the Channel Coding Converse

1. (16 pts) *Fano's inequality without conditioning.* Let $\Pr(X = i) = p_i, i = 1, 2, \ldots, m$ and let $p_1 \geq p_2 \geq p_3 \geq \cdots \geq p_m$. The minimal probability of error predictor of X is $\hat{X} = 1$, with resulting probability of error $P_e = 1 - p_1$.

   (a) (8 pts) Choose $p_2, \ldots, p_m$ so as to maximize $H(X)$ subject to the constraint $1 - p_1 = P_e$ to find an upper bound on $H(X)$ that is a function of the constrain parameter P_e. This is Fano's inequality as expressed in (2.130) without conditioning on $\hat{X}$.

   (b) (8 pts) Now upper bound $H(P_e)$ to provide a lower bound on P_e that corresponds to (2.132) but without the conditioning on Y.

Lecture 10: Differential Entropy

2. (12 pts) *Differential entropy.* Evaluate the differential entropy $h(X) = -\int f \ln f$ for the following:

   (a) The exponential density, $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$.

   (b) The Laplace density, $f(x) = \frac{1}{2}\lambda e^{-\lambda|x|}$.

   (c) The sum of X_1 and X_2, where X_1 and X_2 are independent normal random variables with means μ_i and variances $\sigma_i^2, i = 1, 2$.

3. (24 pts) *Exponential Channel*
   Consider a channel in which the input X is one of two discrete values $X \in \{0, 1\}$. The output Y takes on one of two different distributions depending on the value of X. Specifically,

$$f_{Y|X}(y|x) = \begin{cases} \frac{3e^{-y}}{2} \text{ for } 0 \le y \le \ln 3 \text{ and } 0 \text{ otherwise} & \text{if } x = 0 \\ \frac{3e^{-y}}{2} \text{ for } \ln \frac{3}{2} \le y \le \infty \text{ and } 0 \text{ otherwise} & \text{if } x = 1 \end{cases} \tag{1}$$

   (a) (3 pts) Give the expression for $f_Y(y)$ and show that your expression integrates to 1. *Hints:* Your description of $f_Y(y)$ should have three distinct regions of nonzero density and $f_Y(y) = \sum_x f_{Y|X}(y|x)P_X(x)$.

   (b) (1 pt) Compute $H(X)$ for X equally likely to be 0 or 1.

   (c) (3 pts) Compute $H(X|Y = y)$ for the three cases $0 \le y \le \ln \frac{3}{2}$, $\ln \frac{3}{2} \le y \le \ln 3$, $\ln 3 \le y \le \infty$.

   (d) (3 pts) Compute $H(X|Y)$

   (e) (2 pts) For X equally likely to be 0 or 1, compute $I(X;Y)$.

   (f) (5 pts) Compute $h(Y) = -\int_0^\infty f_Y(y) \ln(f_Y(y))\, dy$. *Note:* Use the natural logarithm ln instead of $\log_2$ to simplify the calculation.

   (g) (5 pts) Compute $h(Y|X)$.

   (h) (2 pts) For X equally likely to be 0 or 1, compute $I(X;Y)$ using $h(Y)$ and $h(Y|X)$

4. (10 pts) *Conditional entropy of a product.*
   This is a question that explores a difference between the entropy H and the differential entropy h.

   (a) (2 pts) For a discrete random variable Y, express $H(aY)$ in terms of $H(Y)$. Assume that $a \neq 0$. Give a simple argument to support your result.

   (b) (3 pts) Find a simplified expression for $H(XY|X)$ involving $H(Y|X)$. Show your derivation. As above, assume that $P(X = 0) = 0$.

   (c) (5 pts) Now consider a continuous random variable Y with pdf $f(Y)$. Find a simplified expression for $h(XY|X)$ involving $h(Y|X)$. Show your derivation. *Hint:* You may use without proof the result $h(aY) = h(Y) + \log|a|$, which we derived in lecture.

5. (9 pts) *Data Processing and Entropy.*

(a) (2 pts) As a brief review, prove that $H(g(X)) \leq H(X)$ for any deterministic function $g(\cdot)$. Here, X is a discrete random variable with a probability mass function.

(b) (2 pts) Show that the inequality of part (a) does not hold for differential entropy by providing a *simple* example where $h(g(X)) > h(X)$. Note that for this part X is a continuous random variable with a probability density function.

(c) (5 pts) Show that $h(g(X)) \leq h(X)$ for the many-to-one mapping $g(\cdot)$ which maps the real line to the interval (-.5, .5] as follows:

$$g(x) = x + n(x), \tag{2}$$

where $n(x)$ is the (possibly negative) integer such that $g(x) \in (-.5, .5]$.

Hint You may use without proof the following inequality, which applies when $f(x+i) \geq 0$ for all x and all i.

$$\sum_{i=-\infty}^{\infty} f(x+i)\log(f(x+i)) \leq \left(\sum_{i=-\infty}^{\infty} f(x+i)\right)\log\left(\sum_{i=-\infty}^{\infty} f(x+i)\right) \tag{3}$$

You may assume that the cumulatives for both X and $g(X)$ are continuous (i.e., there are no mass points in either density function).

6. (10 pts) *More Modulo Mischief*

This problem is a continuation of problem 4.

(a) (2 pts) For positive a and b show that

$$a \log a + b \log b \leq (a+b) \log(a+b). \tag{4}$$

Hint: Showing this fact doesn't require information theory, per se.

(b) (2 pts) Use the generalization of part (a) to prove the hint of problem 4 on problem set 5 as follows: Suppose that $f(x)$ is a probability density function. Prove that

$$\sum_{i=-\infty}^{\infty} f(x+i) \log(f(x+i)) \leq \left(\sum_{i=-\infty}^{\infty} f(x+i) \right) \log \left(\sum_{i=-\infty}^{\infty} f(x+i) \right). \tag{5}$$

(c) (2 pts) For positive a and b show that

$$a \log a + b \log b = (a+b) \log(a+b) + a \log \left(\frac{a}{a+b} \right) + b \log \left(\frac{b}{a+b} \right). \tag{6}$$

(d) (4 pts) Now recall the modulo operation $g(\cdot)$ of problem 4 which maps the real line to the interval (-.5, .5] as follows:

$$g(x) = x + n(x), \tag{7}$$

where $n(x)$ is the unique (possibly negative) integer such that $g(x) \in (-.5, .5]$.

For a continuous random variable X with pdf $f(x)$ define two random variables $Y = g(X)$ and $Z = n(X)$, with $g(x)$ and $n(x)$ as defined above. Note that Y is continuous (with a pdf) and Z is discrete (with a pmf). Show that

$$h(X) = h(Y) + H(Z|Y) \tag{8}$$

Hint: You may use the fact that the conditional pmf for Z given Y is as follows:

$$P(Z = z | Y = y) = \frac{f(y-z)}{\sum_{i=-\infty}^{\infty} f(y-i)} \tag{9}$$

7. (13 pts) *Mutual information for a mixed distribution.(Cong Shen's distribution)* Consider the following channel:

   - The input X is a binary random variable $X \in \{0, 1\}$. For all parts of this problem, assume that X is equally likely to be 0 or 1.
   - The output Y is a neither completely discrete or completely continuous as described below.
   - When the input X equals 0, the output Y is also 0 with probability 1.
   - When the input X equals 1 the output Y is uniformly distributed on the closed interval $\left[\frac{1}{2}, \frac{3}{2}\right]$

   (a) (1 pt) Find $H(X)$.

   (b) (2 pts) Find $H(X|Y)$.

   (c) (6 pts) Ultimately, find the differential entropy $h(Y|X)$. Along the way, you will compute two differential entropies with specific conditioning.

      i. (2 pts) $h(Y|X = 0)$. (Hint: The differential entropy of a discrete random variable is $-\infty$.)

      ii. (2 pts) $h(Y|X = 1)$.

      iii. (2 pts) $h(Y|X)$

   (d) (2 pts) Find $h(Y)$.

   (e) (2 pts) Find $I(X; Y)$.