ECE 5630 - Homework Assignment 5

April 29th 2020

Due to: Tuesday, May 12th, at 2pm.

Instructions: Submission in pairs is allowed. Prove and explain every step in your answers. HW sheets are to be submitted via Canvas.

Assumption: In this homework sheet we assume that alphabets are countable throughout.

1) **Reproving Fano's inequality:** Consider the Markov chain $X \to Y \to \hat{X}$, where X takes values in $\mathcal{X}$ with $|\mathcal{X}| < \infty$, and define the probability of error $P_e := \mathbb{P}(\hat{X} \neq X)$.
   (a) Define the binary random variable $E := \mathbb{1}_{\{\hat{X} \neq X\}}$ and show that $\mathbb{E}[E] = P_e$ and $\mathsf{var}(E) = P_e(1 - P_e)$.
   (b) Using chain rule of entropy to show $H(X|\hat{X}) + H(E|X, \hat{X}) = H(E|\hat{X}) + H(X|E, \hat{X})$.
   (c) Prove the following: $H(X|E, \hat{X}) \leq P_e \log |\mathcal{X}|$.
   (d) Prove $H_b(P_e) + P_e \log |\mathcal{X}| \geq H(X|\hat{X})$, where H_b is the binary entropy function. Explain every step.
   (e) Using the data processing inequality and the results above, show that $P_e \geq \frac{H(X|Y)-1}{\log |\mathcal{X}|}$.

2) **General Fano's inequality:** Let $X, Y \sim P_{X,Y} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$. Define $Q_{XY} = P_X \otimes P_Y$ and let $p := P_{X,Y}(X = Y)$ and $q := Q_{X,Y}(X = Y)$. In words, p and q are the probabilities that $X = Y$ under the joint law P_{XY} or the product of marginals law Q_{XY}, respectively.
   a) Prove that $I(X;Y) \geq p \log \left(\frac{1}{q}\right) - H_b(p)$, where H_b is the binary entropy function.
   Hint: Recall that $I(X;Y) = \mathsf{D}_{\mathsf{KL}}(P_{XY} \| Q_{XY})$ and use f-divergence DPI.
   b) Assume $P_X = \mathsf{Unif}(\mathcal{X})$ (clearly $|\mathcal{X}| < \infty$) and obtain the regular Fano's inequality from Part (a), i.e.,

$$H(X|Y) \leq P_{XY}(X \neq Y) \log |\mathcal{X}| + H_b\big(P_{XY}(X \neq Y)\big).$$

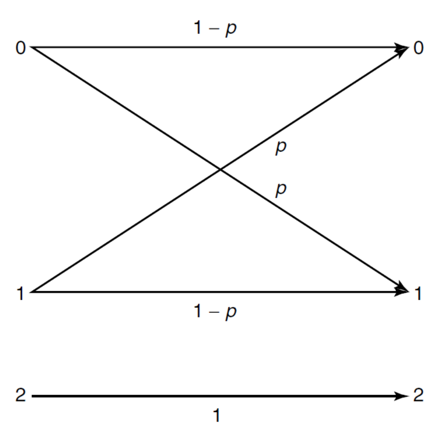
3) **Application of Fano's inequality:** Let $\mathcal{X} = \{1, 2, \ldots, m\}$ and $X \sim P$ with PMF $p(i) = p_i$ for all $i \in \mathcal{X}$ such that $p_1 \geq p_2 \geq \ldots \geq p_m$.
   a) Consider deterministic predictors of the form $\widehat{X} = i$. Find the predictor with minimum probability of error P_e.
   b) Maximize $H(P)$ subject to the constraint that the probability of error of the estimator found in part (a) remains unchanged.
   c) Use part (b) to derive a bound on P_e in terms of the entropy.

4) **Alphabet of noise:** Let $\mathcal{X} = \{0, 1, 2, 3, 4\}$. Consider the channel $Y = X + Z$ where Z is uniformly distributed over $\mathcal{Z} = \{z_1, z_2, z_3\}$. Here, $z_1, z_2, z_3 \in \mathbb{Z}$.
   a) What is the maximum capacity of this channel? Provide a distribution on $\mathcal{X}$ and a choice of alphabet $\mathcal{Z}$ that achieves the capacity.
   b) What is the minimum capacity of this channel? Provide a distribution on $\mathcal{X}$ and a choice of alphabet $\mathcal{Z}$ that achieves the capacity.

5) **Union of channels:** In this problem we want to find the capacity C of the union of two channels. Specifically, let $\mathcal{X}_1 = \{1, \ldots, m\}$ and $\mathcal{X}_2 = \{m+1, \ldots, n\}$ and channels $(\mathcal{X}_1, p_{Y_1|X_1}, \mathcal{Y}_1)$ and $(\mathcal{X}_2, p_{Y_2|X_2}, \mathcal{Y}_2)$, where at each time, one can send a symbol over channel 1 or channel 2 but not both.
   a) Let $X_1 \sim P_1 \in \mathcal{P}(\mathcal{X}_1)$, $X_2 \sim P_2 \in \mathcal{P}(\mathcal{X}_2)$, and

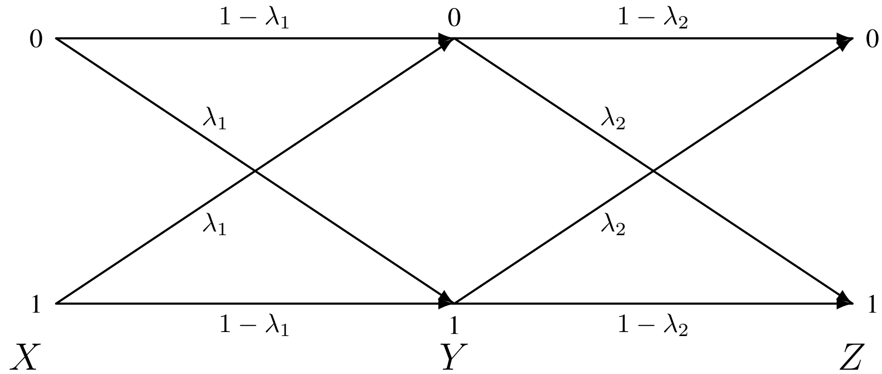
$$X = \begin{cases} X_1, & \text{with probability } p, \\ X_2, & \text{with probability } 1-p. \end{cases}$$

   Find $H(X)$ in terms of $H(X_1)$, $H(X_2)$, and p.
   b) Maximize over p to show that $2^{H(X)} \leq 2^{H(X_1)} + 2^{H(X_1)}$. One can view $2^{H(X)}$ as the effective alphabet size.
   c) Find capacity of the union of two channels C.
   Hint: View 2^C as the effective alphabet size of a channel with capacity C.
   d) Use part (c) to compute the capacity of the following channel.



6) **Preprocessing the output:** Consider a discrete memoryless channel (DMC) $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$.
   a) Prove that it is impossible to strictly increase the channel capacity $\max_{P_X} I(X;Y)$ by preprocessesing the output Y by forming $\tilde{Y} = g(Y)$, giving rise to a new (effective) channel $P_{\tilde{Y}|X}$?
   b) Under what conditions preprocessing does not strictly decrease the capacity?

7) **Cascaded BSCs.** Consider two DMCs $(\mathcal{X}, \mathcal{Y}, P_{Y|X})$ and $(\mathcal{Y}, \mathcal{Z}, P_{Z|Y})$. Let $P_{Y|X}$ and $P_{Z|Y}$ be binary symmetric channels with crossover probabilities λ_1 and λ_2 respectively.



a) What is the capacity C_1 of $P_{Y|X}$?

b) What is the capacity C_2 of $P_{Z|Y}$?

c) We now cascade these channels, obtaining a new effective channel $Q_{Z|X}$ given by $Q_{Z|X}(z|x) = \sum_y P_{Y|X}(y|x)P_{Z|Y}(z|y)$, for all $x, z \in \{0, 1\}$. What is the capacity C_3 of $Q_{Z|X}$? Show $C_3 \leq \min\{C_1, C_2\}$.

d) Now let us actively intervene between channels 1 and 2, rather than passively forwarding the output Y of $P_{Y|X}$ through $P_{Z|Y}$. What is the capacity of channel 1 followed by channel 2 if you are allowed to decode the output Y^n of channel 1 and then re-encode it as $\tilde{Y}^n$ for transmission over channel 2?

Hint: Consider the Markov chain $M \leftrightarrow X^n \leftrightarrow Y^n \leftrightarrow \tilde{Y}^n \leftrightarrow Z^n \leftrightarrow \hat{M}$.)

e) What is the capacity of the cascade in part c) if the end receiver can view *both* Y and Z?

8) **Product channel:** Consider two DMCs $(\mathcal{X}_1, \mathcal{Y}_1, P_{Y_1|X_1})$ and $(\mathcal{X}_2, \mathcal{Y}_2, P_{Y_2|X_2})$ with capacities C_1 and C_2 respectively. A new channel $(\mathcal{X}_1 \times \mathcal{X}_2, \mathcal{Y}_1 \times \mathcal{Y}_2, P_{Y_1,Y_2|X_1,X_2})$, where $P_{Y_1,Y_2|X_1,X_2}(y_1, y_2|x_1, x_2) = P_{Y_1|X_1}(y_1|x_1)P_{Y_2|X_2}(y_2|x_2)$ for all (x_1, x_2, y_1, y_2), is formed. Find the capacity of this channel in terms of C_1 and C_2.

9) **The Z-channel:** This channel has binary input and output alphabets and a channel transition kernel $P_{Y|X}$ as described by the matrix $Q = \left[Q_{x,y}\right]_{x,y\in\{0,1\}}$:

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix},$$

where $Q_{x,y} = P_{Y|X}(y|x)$, for all $x, y \in \{0, 1\}$.

a) Draw a diagram of the Z-channel.

b) Find the capacity of the Z-channel and the maximizing input probability distribution.