EE6340: Information Theory
Problem Set 2

1. *Entropy of a sum.* Let X and Y be random variables that take on values $x_1, x_2, ..., x_r$ and $y_1, y_2, ..., y_s$,respectively. Let $Z = X + Y$.

   (a) Show that $H(Z|X) = H(Y|X)$. Argue that if X, Y are independent,then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$. Thus the addition of *independent* random variables adds uncertainty.

   (b) Give an example (of necessarily dependent random variables) in which $H(X) > H(Z)$ and $H(Y) > H(Z)$.

   (c) Under what conditions does $H(Z) = H(X) + H(Y)$?

2. *Entropy of a disjoint mixture.* Let X_1 and X_2 be discrete random variables drawn according to probability mass functions $p_1(.)$ and $p_2(.)$ over the respective alphabets $\chi_1 = \{1, 2, ..., m\}$ and $\chi_2 = \{m + 1, 2, ..., n\}$.Let

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

   (a) Find $H(X)$ in terms of $H(X_1)$ and $H(X_2)$ and α.

   (b) Maximize over α to show that $2^{H(X)} \leq 2^{H(X_1)} + 2^{H(X_2)}$ and interpret using the notion that $2^{H(X)}$ is the effective alphabet size.

   (c) Let X_1 and X_2 be uniformly distributed over their alphabets.What is the maximizing α and the associated $H(X)$?

3. *Mixing increases entropy.* Show that the entropy of a probability distribution, $(p_1, ...p_i, ...p_j, ...p_m)$, is less than the entropy of the distribution $\left(p_1, ..., \frac{p_i+p_j}{2}, .., \frac{p_i+p_j}{2}, .., p_m\right)$. In general any transfer of probablity that makes the distribution more uniform increases the entropy.

4. *Run length coding.* Let $X_1, X_2, ...X_n$ be (possibly dependent) binary random variables. Suppose one calculates the run lengths $\mathbf{R}$=$(R_1, R_2, ...)$ of this sequence (in order as they occur). For example,the sequence $\mathbf{X} = 0001100100$ yields run lengths $\mathbf{R}$=$(3, 2, 2, 1, 2)$. Compare $H(X_1, X_2, ..., X_n), H(\mathbf{R})$ and $H(X_n, \mathbf{R})$. Show all equations and inequalities, and bound all the differences.

5. *Conditional mutual information vs. unconditional mutual information.* Give examples of joint random variables X, Y and Z such that

   (a) $I(X; Y|Z) < I(X; Y)$,

   (b) $I(X; Y|Z) > I(X; Y)$.

6. *Data processing.* Let $X_1 \to X_2 \to X_3 \to .... \to X_n$ form a Markov chain in this order; i.e., let

$$p(x_1, x_2, ..., x_n) = p(x_1)\, p(x_2|x_1) ...p(x_n|x_{n-1}).$$

   Reduce $I(X_1; X_2, ..., X_n)$ to its simplest form.