Information Theory **Assignment II**

1. Suppose X is a random variable taking values in $\{1, 2, .., n\}$. Let p be a permutation from the set $\{1, 2, .., n\}$ chosen uniformly at random. Show that $H(p(X)) \geq H(X)$. Thus, "shuffling" increases the entropy of a set. □

2. Three random variables X, Y, Z forms a *Markov Chain* if $Pr(Z = z|Y = y, X = x) = Pr(Z = z|Y = y)$. This intutively means that given Y, Z does not depend on X. Confirm the intution by showing that for a Markov chain $Pr(X = x; Z = z|Y = y) = Pr(X = x|Y = y).Pr(Z = z|Y = y)$. Hence conclude that $I(X, Z|Y) = 0$. □

3. Let X, Y, Z be random variables. Define $I(X, Z|Y)$. Show that $I(X, Z|Y) = 0$ if X, Y, Z forms a Markov chain. □

4. If X, Y, Z are random variables, prove that $I(X; YZ) = I(X : Y) + I(X; Z|Y)$. If X, Y, Z forms a Markov chain, expand $I(X; YZ)$ the above equality in two different ways to prove that $I(X, Z) \leq I(X, Y)$. This is called the *Data Processing Inequality* because this tells that processing the received Y "further" at the receiving end to get a "refined" Z does not yield any additional Information. □

5. Suppose random variable X is transmitted and Y received. We would like to estimate X at the receiver applying some function g on X (g can even be a random function) to obtain Z. Show that $Pr(X \neq Z) \geq \frac{H(X|Y)-1}{\log(|X|-1)}$ where $|X|$ denotes the size of the ensemble X. (Note that this is a slightly stronger version of the Fano's inequality proved in the class). □

6. Suppose a noisy discrete communication channel has an input alphabet X and output alphabet Y. Suppose M is the set of messages distributed according to some probability distribution. Suppose M is encoded with coding scheme to produce a codeword in the ensemble X^n, which is transmitted across a noisy communication channel to yield a receiver ensemble Y^n. Show that $I(M; Y^n) \leq \Sigma_{i=1}^{n} I(X_i; Y_i) \leq nC$ where X_i, Y_i indicates the ensembles corresponding to the i^{th} transmission and C the capacity of the channel. Hence show that $H(M) \leq H(M|Y^n) + nC$. □

7. In the above question, Suppose g is an estimater that tries to recover $m \in M$ from $y \in Y^n$ received at the receiver, use Fano bound to dervice a condition on the entropy of the source necessary to ensure that the probability of error in estimation is stricly bounded away from zero irrespective of the value of n. □