Information Theory [MA5103]
Homework 2

1. Let $P = \{p_1, \dots, p_n\}$ and $Q = \{q_1, \dots, q_n\}$ be two probability distributions. Then, we define the relative entropy as follows:

$$D(P\|Q) = \begin{cases} \sum_{i=1}^{n} p_i(\log p_i - \log q_i) & \text{if } q_i = 0 \text{ implies } p_i = 0 \text{ for } \forall i \in \{1, \dots, n\} \\ +\infty & \text{otherwise} \end{cases}$$

   (a) Prove that

$$D(P\|Q) \geq 0$$

   (b) Prove the joint convexity:

$$D(\lambda P + (1-\lambda)P' \| \lambda Q \times (1-\lambda)Q') \leq \lambda D(P\|Q) + (1-\lambda)D(P'\|Q')$$

   Here, $P'' = \{p'_1, \dots, p'_n\}$ and $Q' = \{q'_1, \dots, q'_n\}$ be other two probability distributions.

   (c) By using (a), prove the bound for the entropy: $H(X) \leq \log n$ where X is a random variable with probability P.

   (d) By using (a), prove the positivity of the mutual information $I(X : Y) \geq 0$ where Y is another random variable with probability Q.

   (e) Suppose we have joint probability distributions $P_{X,Y} = \{p(x, y)\}$ and $Q_{X',Y'} = \{q(x, y)\}$ such that $X, X' \in \mathcal{X}$ and $Y, Y' \in \mathcal{Y}$, and define the conditional relative entropy as

$$D(P_{Y|X}\|Q_{Y|X}) = \sum_{x\in\mathcal{X}} p(x) \sum_{y\in\mathcal{Y}} p(y|x)(\log p(y|x) - \log q(y|x))$$

   Then, prove the chain rule:

$$D(P_{X,Y}\|Q_{X,Y}) = D(P_X\|Q_X) + D(P_{Y|X}\|Q_{Y|X})$$

   where $P_X = \{p(x)\}$ and $Q_X = \{q(x)\}$.

2. Suppose we have a random n-bit string $X = (X_1, \dots, X_n)$, where $X_i \in \{0, 1\}$, and Y a random variable. Also, take functions $f_1, \dots, f_n$ from the codomain (target set) of Y to $\{0, 1\}$ and let

$$p_e = \frac{1}{n} \sum_{i=1}^{n} Prob(X_i \neq f_i(Y))$$

then prove the following corollary of Fano's inequality:

$$h(p_e) := H(\{p_e, 1 - p_e\}) \geq \frac{1}{n} H(X|Y)$$

by using other theorems in the lecture if necessarily.

3. Suppose the random variables X, Y, Z with ranges $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ form a Markov chain $X \to Y \to Z$. Show that the mutual information between X and Z is limited by the size of $\mathcal{Y}$:

$$I(X : Z) \leq \log |\mathcal{Y}|$$

4. Let $X_1, X_2, \ldots$ and $X'_1, X'_2, \ldots$ be two Markov chains of random variables with values in a set $\mathcal{X}$ such that they have identical transition probabilities. Then, show that

$$D(P_{X_n} || P_{X'_n}) \geq D(P_{X_{n+1}} || P_{X'_{n+1}})$$

by using 1(e).