Information Theory: Exercise I

1) Let $0 \leq p_1, \ldots, p_m \leq 1$ be such that $\sum_i p_i = 1$. Let $0 \leq q_1, \ldots, q_m \leq 1$ be such that $\sum_i q_i = 1$. Show that

$$-\sum_i p_i \log p_i \leq -\sum_i p_i \log q_i$$

2) Derive and prove a "chain-rule" expression for $I((X_1, \ldots, X_n); Y)$

3) What is larger, $I(X;Y|Z)$ or $I(X;Y)$?

4) What is larger, $H(X|Y)$ or $H(f(X)|Y)$?

5) What is larger, $H(X|Y)$ or $H(X|g(Y))$?

6) What is larger, $H(X|Y)$ or $H(f(X,Y)|Y)$?

7) What is larger, $H(X|Y)$ or $H(X|g(X,Y))$?

8) Fano's inequality: Let $X \in \{1, \ldots, m\}$ be a random variable and let Y be any random variable. Assume that for some function g to the range $\{1, \ldots, m\}$ and for some $\epsilon < 1/2$ we have $\text{Prob}[X = g(Y)] \geq 1 - \epsilon$. Prove that

$$H(X|Y) \leq H(\epsilon, 1-\epsilon) + \epsilon \cdot \log(m-1)$$

9) For a (finite) set A of random variables, denote by $H(A)$ the joint entropy of all the variables in A. Prove that for any two (finite) sets A, B of random variables

$$H(A \cup B) + H(A \cap B) \leq H(A) + H(B)$$