Review of Basic Probabilities

1. Let $X \in \{x_1, x_2, x_3\}$ be a discrete random variable with probability mass function $p(x_1) = 0.1$, $p(x_2) = 0.3$, and $p(x_3) = 0.6$. 
   
   (a) What is the probability that $p(X) \in (0.01, 0.5)$? 
   (b) What is $E[e^{p(X)}]$? 
   (c) What is $E[-\log p(X)]$? 
   (d) What is $E[1/p(X)]$? 
   (e) Recall that I said in the class, the SIC of a rare event is large, and a frequent event is small. Now, consider the function $y(X) = 1/p(X)$ which satisfies our the intuition of SIC. Discuss some of the drawbacks for using this definition for SIC?

Understanding Entropy

2. The p.m.f of a discrete random variable $X$ over a set of alphabet $A = \{1, 2, 3, 4\}$ is $\{p_1, p_2, p_3, p_4\}$. The p.m.f of random variable $Y$ over the set $A_1 = \{1, 2\}$ is $\{\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\}$, and the p.m.f of random variable $Z$ over set $A_2 = \{3, 4\}$ is $\{\frac{p_3}{p_3 + p_4}, \frac{p_4}{p_3 + p_4}\}$. Use the definition of entropy to show that 

$$H(X) = H(p_1 + p_2, p_3 + p_4) + (p_1 + p_2)H(Y) + (p_3 + p_4)H(Z).$$

Convince yourself that it does not matter how you partition the original sample space of $A$ into two smaller groups $A_1$ and $A_2$, it is always true that $H(X) = H(p, 1 - p) + pH(Y) + (1 - p)H(Z)$ where $p$ and $1 - p$ are the statistical weights for each smaller group. A way to view entropy is to use a example of finding a hidden treasure (one of the elements of $A$). $H(p, 1 - p)$ is the uncertainty involved in whether the treasure is hidden in $A_1$ or $A_2$. If the treasure is in $A_1$ (with probability $p$), the additional uncertainty is $H(Y)$. If the treasure is in $A_2$ (with probability $1 - p$), the additional uncertainty is $H(Z)$. Note that $\{\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\}$ used in $H(Y)$ is the conditional probability given that the treasure is in group $A_1$. Similarly for $A_2$, this conditional probability is $\{\frac{p_3}{p_3 + p_4}, \frac{p_4}{p_3 + p_4}\}$. Intuitively, the total uncertainty (information) to get to the treasure should not change with how you partition the group. So $H(X) = H(p_1 + p_2, p_3 + p_4) + (p_1 + p_2)H(Y) + (p_3 + p_4)H(Z)$.

3. Let $X_1 \in [0, 1]$ and $X_2 \in [0, 1]$ be two independent random variables with $X_1 \sim Bern(2/3)$ and $X_2 \sim Bern(1/2)$. Let $Y = \min\{X_1, X_2\}$, $Z = \max\{X_1, X_2\}$.
   
   (a) Compute $H(Y), H(Y, Z), H(Y|Z)$, and $I(Z; Y)$.
   (b) Based on (a), if you know $Z$, does that help you to guess $Y$ more accurately?

4. The sum of the faces of two normal dice when thrown is 7. How much information does this fact supply us with? Explain your answer. Note: outcomes such as $(6, 1)$ and $(1, 6)$ are to be considered as being different. You should view this as follows. In the beginning, you don’t know anything, so how much information (uncertainty) is there in the outcomes. Next, after you see the outcome (sum = 7), how much uncertainty is remained? Then the information gain, i.e., information supplied to us, is the original amount of information subtract the remaining uncertainty.

5. Let $X$ be a discrete random variable with probability mass function $p(X) = \{p_0, p_1, \ldots, p_m\}$. Let $Y$ be another discrete random variable with probability mass function $q(Y) = \{q_0, q_1, \ldots, q_m\}$, where 

$$q_i = p_i \quad i = 0, 1, \ldots, j - 2, j + 1, \ldots, m$$
\[ q_j = q_{j-1} = \frac{p_j + p_{j-1}}{2} \]

How is \( H(Y) \) related to \( H(X) \) (greater, equal, or less)? Prove your answer.