Assignment 1

1. **Computing entropy.** (2 points each) In this problem, you are to compute the entropy of various probability distributions, represented by sets of numbers in brackets. If there are \( K \) numbers in brackets, then there are \( K \) possible values of the corresponding random variable, whose probabilities are given by the numbers. In other words, a fair die would be written \( \left[ \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right] \). Compute the entropy of the following distributions (don’t use a calculator unless the problem says you can):

   a \( \left[ \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right] \)
   b \( \left[ \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0, 0, 0 \right] \)
   c Compare answers a and b. Make a general statement about entropy calculations.
   d \( \left[ \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{9}{4} \right] \) (you can use a calculator for this one)
   e Compare answers a and d. Can you offer a conjecture based upon these? (No need to prove it right now.)
   f \( \left[ \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8} \right] \)
   g \( \left[ \frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n} \right] \)

2. A die is labeled from 1 to 6. Assuming it’s fair, what is the entropy of the roll (you can leave a “log” in your answer)? The die is relabeled with the even numbers from 2 to 12. What is its entropy now? (2 points)

3. **Independence and entropy.** (A) Give the entropy, in bits, of four fair, 8-sided dice. (B). Suppose you drill a hole in each die, and tie them all together with a string. Will the entropy of the dice be higher or lower than the answer from part (A)? Why? (10 points).

4. **Coin Flips.** (From Cover and Thomas, 2nd edition, Problem 2.1, part (a)). (10 points). A fair coin is flipped until the first head occurs. Let \( X \) denote the number of flips required. Find the entropy \( H(X) \) in bits. The following expressions may be useful:

\[
\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}, \quad \sum_{n=0}^{\infty} n r^n = \frac{r}{(1-r)^2}
\]

5. **Entropy of functions.** (From Cover and Thomas, 2nd edition, Problem 2.2). (5 points). Let \( X \) be a random variable taking on a finite number of values. What is the (general) inequality relationship of \( H(X) \) and \( H(Y) \) if

(a) \( Y = 2^X \)
6. Minimum entropy. (From Cover and Thomas, 2nd edition, Problem 2.3). (5 points). What is the minimum value of \( H(p_1, \ldots, p_n) = H(p) \) as \( p \) ranges over the set of \( n \)-dimensional probability vectors? Find all \( p \)'s that achieve this minimum.

7. Entropy of functions of a random variable. (From Cover and Thomas, 2nd edition, Problem 2.4). (5 points). Let \( X \) be a discrete random variable. Show that the entropy of a function of \( X \) is less than or equal to the entropy of \( X \) by justifying the following steps:

\[
H(X, g(X)) \overset{(a)}{=} H(X) + H(g(X) | X) \\
\overset{(b)}{=} H(X), \\
H(X, g(X)) \overset{(c)}{=} H(g(X)) + H(X | g(X)) \\
\overset{(d)}{\geq} H(g(X)).
\]

Thus, \( H(g(X)) \leq H(X) \).

8. Conditional mutual information vs. unconditional mutual information. (From Cover and Thomas, 2nd edition, Problem 2.6). (5 points). Give examples of joint random variables \( X, Y \), and \( Z \) such that

(a) \( I(X; Y | Z) < I(X; Y) \).

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

9. Example of joint entropy. (From Cover and Thomas, 2nd edition, Problem 2.12, parts (a)-(e)). (2 points each).

Let \( p(x, y) \) be given by

<table>
<thead>
<tr>
<th></th>
<th>Y 0</th>
<th>Y 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>X 0</td>
<td>1/3</td>
<td>1/3</td>
</tr>
<tr>
<td>X 1</td>
<td>0</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Find:

(a) \( H(X), H(Y) \).

(b) \( H(X|Y), H(Y|X) \).

(c) \( H(X,Y) \).

(d) \( H(Y) - H(Y|X) \).

(e) \( I(X; Y) \).
10. Relative entropy. (From Cover and Thomas, 2nd edition, Problem 2.37). (10 points). Let $X, Y, Z$ be three random variables with a joint probability mass function $p(x, y, z)$. The relative entropy between the joint distribution and the product of the marginals is the following:

$$D(p(x, y, z)||p(x)p(y)p(z)) = E \left[ \log \frac{p(x, y, z)}{p(x)p(y)p(z)} \right]$$

Expand this in terms of entropies. When is this quantity zero? Justify your answer in detail.

   Let $X, Y, Z$ be three binary Bernoulli($\frac{1}{2}$) random variables that are pairwise independent; that is, $I(X; Y) = I(X; Z) = I(Y; Z) = 0$.
   (a) Under this constraint, what is the minimum value for $H(X, Y, Z)$?
   (b) Give an example achieving this minimum.

   For extra credit (up to 10 additional points), prove that your answer is one of the best possible answers.

12. Mutual information of heads and tails. (From Cover and Thomas, 2nd edition, Problem 2.43). (5 points).
   (a) Consider a fair coin flip. What is the mutual information between the top and bottom sides of the coin?
   (b) A six-sided fair die is rolled. What is the mutual information between the top side and the front face (the side most facing you)?