1. **Entropy of Hamming Code.**
   Consider information bits $X_1, X_2, X_3, X_4 \in \{0, 1\}$ chosen uniformly at random, together with check bits $X_5, X_6, X_7$ chosen to make the parity of the circles even.

![Diagram of Hamming Code]

Thus, for example,

![Diagram of Example 1]

becomes

![Diagram of Example 2]

That is, 1011 becomes 1011010.

(a) What is the entropy of $H(X_1, X_2, ..., X_7)$?
Now we make an error (or not) in one of the bits (or none). Let \( Y = X \oplus e \), where \( e \) is equally likely to be \((1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 0, 1), \) or \((0, 0, \ldots, 0)\), and \( e \) is independent of \( X \).

(b) What is the entropy of \( Y \)?

(c) What is \( H(X|Y) \)?

(d) What is \( I(X; Y) \)?

2. **Entropy of functions of a random variable.**

Let \( X \) be a discrete random variable.

(a) 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 \( H(g(X)) \leq H(X) \).

(b) A Data-Processing Inequality: Show that if \( Z = g(Y) \) then \( H(X|Y) \leq H(X|Z) \).

3. **Maximum and minimum entropy**

(a) We have three fair dice, each with six sides. You can put any integer on each side of each die. (You are allowed to put the same integers multiple times.) We toss the dice and record the sum of the three numbers on top.

(i) What is the maximum entropy of the sum?

(ii) What is the minimum entropy of the sum?

(b) Repeat part (a) for \( k \) dice.

4. **Entropy of time to first success.**

A fair coin is flipped until the first head occurs. Let \( X \) denote the number of flips required.
(a) Find the entropy \( H(X) \) in bits. The following expressions may be useful:
\[
\sum_{n=1}^{\infty} r^n = r/(1 - r), \quad \sum_{n=1}^{\infty} nr^n = r/(1 - r)^2.
\]

(b) Find an “efficient” sequence of yes-no questions of the form, “Is \( X \) contained in the set \( S \)?”. Compare \( H(X) \) to the expected number of questions required to determine \( X \).

(c) Let \( Y \) denote the number of flips until the second head appears. Thus, for example, \( Y = 5 \) if the second head appears on the 5th flip. Argue that \( H(Y) = H(X_1 + X_2) < H(X_1, X_2) = 2H(X) \), and interpret in words.

5. **Example of joint entropy.**

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

<table>
<thead>
<tr>
<th>( x )</th>
<th>( 0 )</th>
<th>( 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0 )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{4} )</td>
</tr>
<tr>
<td>( 1 )</td>
<td>( 0 )</td>
<td>( \frac{1}{2} )</td>
</tr>
</tbody>
</table>

Find

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

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

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

(d) \( I(X; Y) \).

6. **Infinite entropy.** This problem shows that the entropy of a discrete random variable can be infinite. Let \( A = \sum_{n=2}^{\infty} (n \log^2 n)^{-1} \). (It is easy to show that \( A \) is finite by bounding the infinite sum by the integral of \((x \log^2 x)^{-1}\).) Show that the integer-valued random variable \( X \) distributed as \( P(X = n) = (An \log^2 n)^{-1} \) for \( n = 2, 3, \ldots \) has \( H(X) = +\infty \).

7. **A measure of correlation.**

Let \( X_1 \) and \( X_2 \) be identically distributed with positive entropy, but not necessarily independent. Note that \( H(X_1) = H(X_2) \). Let

\[
\rho = 1 - \frac{H(X_2|X_1)}{H(X_1)}.
\]
(a) Show that $0 \leq \rho \leq 1$
(b) Show that $I(X_1; X_2) = \rho H(X_1)$
(c) Show that $\rho = 0$ iff $X_1$ is independent of $X_2$
(d) Show that $\rho = 1$ iff there exists a one-to-one function $g$ such that $X_1 = g(X_2)$ with probability one.

8. Two looks.
Here is a statement about pairwise independence and joint independence. Let $X, Y_1,$ and $Y_2$ be binary random variables. If $I(X; Y_1) = 0$ and $I(X; Y_2) = 0,$ does it follow that $I(X; Y_1, Y_2) = 0$?

(a) Yes or no? Prove or provide a counterexample.
(b) If $I(X; Y_1) = 0$ and $I(X; Y_2) = 0$ in the above problem, does it follow that $I(Y_1; Y_2) = 0$?

9. Markov’s inequality for probabilities.
Let $p(x)$ be a probability mass function. Prove, for all $d \geq 0,$

$$P(p(X) \leq d) \log \left(\frac{1}{d}\right) \leq H(X).$$

10. Smallest Typical Set
We have a memoryless source $U,$ i.e., $U_1, U_2, \ldots,$ are i.i.d. $\sim U,$ where $U$ takes values in the finite alphabet $\mathcal{U}$. Let $u^n$ denote the $n$-tuple $(u_1, u_2, \ldots, u_n)$ and $p(u^n)$ be its probability, i.e.,

$$p(u^n) = \Pi_{i=1}^{n} P_U(u_i)$$

Let $\delta > 0$ and for every $n$ let $B^{(n)} \subseteq \mathcal{U}^n$ be an arbitrary set of source sequences satisfying $|B^{(n)}| \leq 2^{n(H(U) - \delta)}.$ Prove that:

$$\lim_{n \to \infty} \Pr(U^n \in B^{(n)}) = 0.$$

In words, the typical set $A^{(n)}_c$ (defined in class) is essentially smallest (on an exponential scale) among the sets that have non-negligible probability.