Solution to Homework #1

1. **Entropy of functions of a random variable.** 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 (a), (b), (c) and (d):

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

Hence $H(g(X)) \leq H(X)$.

**Solution:**

(a) By the chain rule for entropies.

(b) Given $X$, $g(X)$ has a fixed value. Hence

\[
H(g(X)|X) = \sum_x p(x)H(g(X)|X = x) = \sum_x 0 = 0.
\]

(c) By the chain rule for entropies.

(d) Follows because the (conditional) entropy of a discrete random variable is nonnegative, i.e., $H(X|g(X)) \geq 0$, with equality iff $g(X)$ is a one-to-one mapping.

2. **A measure of correlation.** Let $X_1$ and $X_2$ be identically distributed, but not necessarily independent. Let

\[
\rho = 1 - \frac{H(X_1|X_2)}{H(X_1)}.
\]

(a) Show $\rho = \frac{I(X_1;X_2)}{H(X_1)}$.

(b) Show $0 \leq \rho \leq 1$.

(c) When is $\rho = 0$?

(d) When is $\rho = 1$?

**Solution:**

(a) $\rho = \frac{H(X_1) - H(X_1|X_2)}{H(X_1)} = \frac{I(X_1;X_2)}{H(X_1)}$

(b) $0 \leq \rho \leq 1$ follows easily because $0 \leq H(X_1|X_2) \leq H(X_1)$. 

1
(c) \( \rho = 0 \) iff \( I(X_1; X_2) = 0 \), i.e., \( X_1 \) and \( X_2 \) are independent.
(d) \( \rho = 1 \) iff \( H(X_1|X_2) = 0 \), i.e., \( X_1 \) is a function of \( X_2 \).

3. Example of joint entropy. Let \( p(x, y) \) be given by

\[
\begin{array}{c|cc}
X & 0 & 1 \\
\hline
Y & \frac{1}{3} & \frac{1}{3} \\
0 & \frac{1}{3} & \frac{1}{3} \\
1 & 0 & \frac{1}{3}
\end{array}
\]

Find
(a) \( H(X) \), \( H(Y) \).
(b) \( H(X|Y) \), \( H(Y|X) \).
(c) \( H(X, Y) \).
(d) \( H(Y) - H(Y|X) \).
(e) \( I(X; Y) \).
(f) Draw a Venn diagram for the quantities in (a) through (e).

Solution:
(a) \( H(X) = \frac{4}{3} \log \frac{3}{2} + \frac{1}{3} \log 3 = \log 3 - \frac{2}{3} = 0.918 \) bits = \( H(Y) \).
(b) \( H(X|Y) = \frac{1}{3}H(X|Y = 0) + \frac{2}{3}H(X|Y = 1) = \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3} = 0.667 \) bits = \( H(Y|X) \).
(c) \( H(X, Y) = 3 \times \frac{1}{3} \log 3 = \log 3 = 1.585 \) bits. Alternatively, \( H(X, Y) = H(X) + H(Y|X) = 1.585 \) bits.
(d) \( H(Y) - H(Y|X) = \log 3 - \frac{2}{3} - \frac{2}{3} = \log 3 - \frac{4}{3} = 0.251 \) bits.
(e) \( I(X; Y) = H(Y) - H(Y|X) = 0.251 \) bits.
(f)

4. Mixing increases entropy. Show that the entropy of the probability distribution

\( (p_1, \ldots, p_i, \ldots, p_j, \ldots, p_m) \)
is less than the entropy of the distribution

\[(p_1, \ldots, \frac{p_i + p_j}{2}, \ldots, \frac{p_i + p_j}{2}, \ldots, p_m).\]

(hint: \(\log_e x \leq x - 1\) for \(x > 0\).)

**Solution:**

Let \(P_1 = (p_1, \ldots, p_i, \ldots, p_j, \ldots, p_m)\) and \(P_2 = (p_1, \ldots, \frac{p_i + p_j}{2}, \ldots, \frac{p_i + p_j}{2}, \ldots, p_m)\).

\[
H_e(P_1) - H_e(P_2) = -p_i \ln p_i - p_j \ln p_j + 2 \frac{p_i + p_j}{2} \ln \frac{p_i + p_j}{2} \\
= p_i \ln \frac{p_i + p_j}{2p_i} + p_j \ln \frac{p_i + p_j}{2p_j} \\
\leq p_i \left(\frac{p_i + p_j}{2p_i} - 1\right) + p_j \left(\frac{p_i + p_j}{2p_j} - 1\right) \\
= 0.
\]

5. **Relative entropy is not symmetric.** Let the random variable \(X\) have three possible outcomes \{a, b, c\}. Consider two distributions \(p(x)\) and \(q(x)\) on this random variable.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>(p(x))</th>
<th>(q(x))</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1/2</td>
<td>1/3</td>
</tr>
<tr>
<td>b</td>
<td>1/4</td>
<td>1/3</td>
</tr>
<tr>
<td>c</td>
<td>1/4</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Calculate \(H(p), H(q), D(p\|q)\) and \(D(q\|p)\). Verify that in this case \(D(p\|q) \neq D(q\|p)\).

**Solution:**

\[
H(p) = \frac{1}{2} \log 2 + \frac{1}{4} \log 4 + \frac{1}{4} \log 4 = 1.5 \text{ bits}
\]

\[
H(q) = 3 \times \frac{1}{3} \log 3 = 1.58496 \text{ bits}
\]

\[
D(p\|q) = \frac{1}{2} \log \frac{3}{2} + \frac{1}{4} \log \frac{4}{3} + \frac{1}{4} \log \frac{4}{3} = \log 3 - 1.5 = 0.08496 \text{ bits}
\]

\[
D(q\|p) = \frac{1}{3} \log \frac{3}{2} + \frac{1}{3} \log \frac{4}{3} + \frac{1}{3} \log \frac{4}{3} = -\log 3 + \frac{5}{3} = 0.0817 \text{ bits}
\]

It is clear that \(D(p\|q) \neq D(q\|p)\).

6. **Discrete entropies.** Let \(X\) and \(Y\) be two independent integer-valued random variables. Let \(X\) be uniformly distributed over \(\{1, 2, \ldots, 8\}\), and let \(Pr\{Y = k\} = 2^{-k}\), for \(k = 1, 2, 3, \ldots\)

(a) Find \(H(X)\).
(b) Find $H(Y)$.

(c) Find $H(X + Y, X - Y)$.

*hint:* For (b), the following expression may be useful
\[
\sum_{n=0}^{\infty} nr^n = \frac{r}{(1 - r)^2}.
\]

For (c), if you do not find a direct way, try to use
\[
I(X, Y; X + Y, X - Y)
= H(X, Y) - H(X, Y | X + Y, X - Y)
= H(X + Y, X - Y) - H(X + Y, X - Y | X, Y).
\]

**Solution:**

(a) $H(X) = \log 8 = 3$ bits.

(b) $H(Y) = \sum_k 2^{-k} \log 2^k = \sum_k k 2^{-k} = \frac{1/2}{(1-1/2)^2} = 2$ bits.

(c) Since $(X, Y) \rightarrow (X + Y, X - Y)$ is a one-to-one mapping,
\[
H(X + Y, X - Y) = H(X, Y) = H(X) + H(Y) = 5 \text{ bits.}
\]

Alternatively, use the hint. It is clear $H(X, Y | X + Y, X - Y) = 0$ and $H(X + Y, X - Y | X, Y) = 0$. Hence $H(X + Y, X - Y) = H(X, Y) = H(X) + H(Y) = 5$ bits.