1. Consider

a. a continuous source with the following probability density function. Construct an optimal 2 bit/sample scalar quantizer for this source.

\[
\begin{align*}
\text{Figure 1: The probability density function for source } X
\end{align*}
\]

By symmetry, it is sufficient to consider

one bit quantization of (a)

From the class, we know that

\[
\begin{align*}
b_j &= \frac{a_{j-1} + a_j}{2} \\
\bar{a}_j &= \frac{\int_{b_{j-1}}^{b_j} x f_X(x) dx}{\int_{b_{j-1}}^{b_j} f_X(x) dx}
\end{align*}
\]

Since

\[
\begin{align*}
a_j &= \frac{\int_{-1}^{b_j} x (1+x) dx}{\int_{-1}^{1} (1+x) dx} = \frac{\frac{1}{3} b_j^3 + \frac{i}{2} b_j^2 + \frac{i}{2}}{\frac{2}{3} b_j^3 + \frac{i}{2} b_j^2 - \frac{i}{2} + 1} \\
\bar{a}_j &= \frac{-\frac{1}{3} b_j^3 + \frac{1}{2} b_j^2 - b_j}{-\frac{1}{2} b_j^2 - b_j}
\end{align*}
\]

Thus,

\[
\begin{align*}
b_j &= \frac{a_{j-1} + a_j}{2} \implies a_j = -0.888 \quad \bar{a}_j = -0.17
\end{align*}
\]

Therefore, the 2 bit quantization points are

\[
\begin{bmatrix}
-2.83 & -1.42 & -0.88 & -0.17
\end{bmatrix}
\]
b. an iid source which produces symbols from alphabet \( \{ A, B, C \} \) with probabilities \( \{0.7, 0.1, 0.2\} \). Construct a ternary Tunstall code of length 2 and compute its efficiency.

\[
\text{Average length of source phrase} = 1 + 0.7 + 0.4 \cdot 3 + 0.3 + 3 = 2.5333.
\]

\[
\text{Average number of code symbols/source symbol} = \frac{2}{2.5333} = 0.7876.
\]

\[
\kappa = \frac{H_3(5)}{0.7896} = \frac{0.7298}{0.7896} = 92.43\%.
\]
2. Consider an i.i.d binary source, \( S_0 \), where the probability of a 0 is equal to \( p \). If \( p \) is close to 1, the source will emit long strings of 0s, separated by 1s. A typical output sequence of this source would be:

\[
00100001010110000101\ldots
\]

Consider a new source \( S \) with symbols \( t_0, t_1, t_2, t_3, t_4, \ldots \), which emits the symbol \( t_i \) when the source \( S_0 \) emits a run of \( i \) 0s followed by a 1 (note that this source symbols have an infinite alphabet). That is for the above binary output of the source \( S_0 \), the source \( S \) emits the sequence:

\[
t_2t_3t_1t_0t_4t_1\ldots
\]

(a) Find the entropy (base 2) of the source \( S \). Do not leave your answer in terms of an infinite series but rather evaluate the infinite sum. The answer should be a function of \( p \).

\[
\begin{align*}
\text{Let } & \quad p(\leq z \leq r) = p^r (1-p) \quad \Delta \quad p^r, \quad i = 0, 1, \\
\text{thus } & \quad H(S) = - \sum_{i=0}^{\infty} p^r \log_2 p^r \\
& = - \sum_{i=0}^{\infty} p^r (1-p) i \log_2 (p^r (1-p)) \\
& = - \left( \sum_{i=0}^{\infty} \frac{i}{p^r} (1-p) \right) (1-p) \log_2 (1-p) - \left( \sum_{i=0}^{\infty} \frac{i}{p^r} \right) (1-p) \log_2 (1-p) \\
& = - \frac{p}{1-p} (1-p) \log_2 (1-p) - \frac{i}{1-p} (1-p) \log_2 (1-p) \\
& = - \frac{p}{1-p} \log_2 (1-p) - \log_2 (1-p)
\end{align*}
\]

(15 points)

(b) Calculate the average number of source symbols from source \( S_0 \) in a symbol from source \( S \).

\[
\text{It is the same as computing the average length of the } i \text{ th source symbol,} \\
\sum_{i=0}^{\infty} (i+1) \frac{i}{p^r} = \sum_{i=0}^{\infty} \frac{i}{p^r} + 1 = \frac{1}{(1-p)} \frac{p}{1-p} + 1 = \frac{p}{1-p} + 1 = \frac{i}{1-p}
\]

(20 points)
3. Consider the following two channels in parallel:

![Figure 2: Two channels used in parallel.](image)

a. Consider $P_X(.)$ be a pmf on $X$, such that $\sum_{x \in U} P_X(x) = 1 - \sum_{x \in Z} P_X(x)$. Show that

$$H(Y|X) = \alpha H(V|U) + (1 - \alpha) H(W|Z)$$

(21 points)

Define a new random variable $E = \{ X \in U \text{ with probability } \alpha \}
\{ X \in Z \text{ with probability } 1 - \alpha \}$

Thus

$$H(\gamma, E|X) = H(\gamma|X) + H(E|\gamma, X)$$

$$= H(E|X) + H(\gamma|X, E)$$

Both $H(E|Y, X)$ and $H(E|X)$ are 0 because $E$ is a function of $X$.

Thus

$$H(\gamma|X) = H(\gamma|X, E)$$

$$= \alpha H(\gamma|X, \{ X \in U \}) + (1 - \alpha) H(\gamma|X, \{ X \in Z \})$$

$$= \alpha H(V|U) + (1 - \alpha) H(W|Z)$$

b. Show that there is a fixed number $\alpha^*$, $0 \leq \alpha^* \leq 1$ such that

$$C_{\text{parallel}} = \alpha^* C_1 + (1 - \alpha^*) C_2 + h(\alpha^*),$$

where $C_1, C_2$ are the capacity of channels associated with $P(V|U)$ and $P(W|Z)$, $h(\alpha) = \alpha \log_2 \frac{1}{\alpha} + (1 - \alpha) \log_2 \frac{1}{1 - \alpha}$.

(27 points)

Since it is easy to see that

$$H(\gamma) = h(\alpha) + \alpha H(V) + (1 - \alpha) H(W),$$

we have

$$I(X: Y) = \alpha I(V: U) + (1 - \alpha) I(W: Z) + h(\alpha)$$

Therefore

$$C_{\text{parallel}} = \max_{P_X} I(X: Y) = \alpha^* C_1 + (1 - \alpha^*) C_2 + h(\alpha^*)$$
From part b, we have \[ C = \alpha^* h(p) + (1-\alpha^*) h(\bar{y}) + h(\alpha^*) \]

\( \alpha^* \) can be obtained as follows, let \( f(x) = x h(p) + (1-x) h(\bar{y}) + h(x) \)

\[ f'(x) = h(p) - h(\bar{y}) + h'(x) \]

\[ f'(x) = 0 \implies \alpha^* = (h')^{-1}(h(\bar{y}) - h(p)) \]

Therefore \[ C = \alpha^* h(p) + (1-\alpha^*) h(\bar{y}) + h(\alpha^*) \]

with \( \alpha^* = (h')^{-1}(h(\bar{y}) - h(p)) \)

c. Consider a discrete memoryless channel with 4 input and 4 output symbols \{0, 1, 2, 3\} such that

\[
\begin{align*}
P(0|0) &= P(1|1) = 1 - p \\
P(1|0) &= P(0|1) = p \\
P(2|2) &= P(3|3) = 1 - q \\
P(3|2) &= P(2|3) = q,
\end{align*}
\]

where \( 0 \leq p \leq q \leq 1 \) are known and fixed scalars. Compute the capacity of this channel.

**Hint:** your answer should be in terms of \( p, q, h(\cdot), \) and \( (h')^{-1}(\cdot) \).

(17 points)
4. Assume that an overall parity check is appended to the Golay (23,12) code

\[
H = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{bmatrix}
\]

to make a (24,12) code.

a. Give the parity check matrix of the code with a unit matrix on the right.

b. Give the generator matrix of the code.

c. What is the minimum distance of the (24,12) code? Find a code word whose Hamming weight is equal to the minimum distance of the code?

d. What is the guaranteed erasure correction capability of the code?

e. What is the MAXIMUM number of erasures that can be corrected by the code?

\[ \text{Let} \quad H = \begin{bmatrix} A & I \end{bmatrix}_{12 \times 24} \]

The parity check matrix for (12, 12) code is

\[ H' = \begin{bmatrix} A_{12 \times 12} \enspace I_{12 \times 12} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} B \enspace I \end{bmatrix}_{12 \times 24} \]

\[ G' = \begin{bmatrix} I \enspace B^T \end{bmatrix} \]

The minimum distance is 8

\[ \begin{bmatrix} 1 
1 
1 
1 
1 
1 
1 
1 
1 
1 \end{bmatrix} \]

is a (24,12) code word with Hamming weight 8

The guaranteed erasure correction capability of the code is 7.

The maximum number of erasures that can be corrected is 8.