1. **Fano Inequality Take Two:**

We want to estimate a quantity \( X \) taking values in a set \( \mathcal{X} \) via the observation \( Y \). Traditionally, a good standard estimator \( \hat{X}(Y) \) is a function of \( Y \), such that \( \hat{X}(Y) = X \) with high probability. In this problem, instead of producing a single value, the estimator produces a list of values in \( \mathcal{X} \), say \( L(Y) \) which depends on the observation \( Y \). The number of elements in the list \( |L| \) is assumed to be constant. We define the probability of error as the probability that the real quantity \( X \) is not in the list \( L(Y) \), i.e. \( P_e = P(X \notin L(Y)) \). Show that

\[
H(X|Y) \leq P_e \log |X| + (1 - P_e) \log |L| + H(P_e)
\]  

(1)

2. You want to communicate with an exploration crew member who is currently lost somewhere on the surface of Mars. Due to the special layers of gases existed on Mars, whatever signal that is transmitted from your spaceship to the surface of Mars, it is being amplified in a peculiar way. Specifically, if the transmitted signal is \( X \) then the received signal on the Mars surface is \( Y = X^2 \). Compounding to this problem, you only have the equipment that transmits only symbols \( X \in \{-1, 1\} \).

(a) What is the maximum rate that you can transmit information reliably to the Mars surface?

(b) To increase the maximum rate, an engineer on board notices that using a special carrier frequency, the layer of gases would scatter the transmitted signal in such a way that the received signal is \( Y = X^2 + Z \) where \( Z = \{0, 1\} \) is independent random noise with \( P(Z = 1) = P(Z = 0) = 1/2 \). He thought that transmitting signals at this carrier frequency would allow us to increase the transmission rate. Do you agree? Explain why or why not?

(c) Being the most clever person on the spaceship (at least you thought so, because none of your exploration team members has taken the information theory class), you test other carrier frequencies, and realize that for one particular frequency, the received signal is \( Y = X^2 + XZ \) where \( Z \) is described in (b). What is the maximum transmission rate that you can achieve using this carrier frequency?

(d) Intuitively, explain why your approach (c) is better/worse/same as that of other engineer in (b)?

3. Determine the capacity of the channel with following transition probability matrix:

\[
Q = \begin{bmatrix}
1 - p & p & 0 \\
0 & p & 1 - p \\
\end{bmatrix}
\]

4. A binary communication channel is made of two cascading subchannels: The transition probability matrix of the first and second subchannels are:

\[
Q_{Y|X} = \begin{bmatrix}
\frac{3}{4} & \frac{1}{4} \\
\frac{1}{4} & \frac{3}{4} \\
\end{bmatrix}
\] and \( Q_{Z|Y} = \begin{bmatrix}
\frac{2}{5} & \frac{3}{5} \\
\frac{3}{5} & \frac{2}{5} \\
\end{bmatrix} \), respectively.

(a) Determine the capacity \( C_1 \) of first subchannel

(b) Determine the capacity \( C_2 \) of the second subchannel

(c) Determine the capacity \( C \) of the overall combined channel

(d) Do you expect that \( C \) will be larger than \( C_1 \) or smaller? Why?

5. Determine the capacity of the channel \( Y = X \mod 100 \) where the input \( X = \{0, 1, 2, \ldots, 200\} \).

6. Determine the capacity of the channel \( Y = (X + Z) \mod 2 \) where the input \( X = \{0, 1, 2\} \), \( Z = \{0, 1\} \), \( Z \) is i.i.d with \( P(Z = 0) = P(Z = 1) = 1/2 \).
7. Three different symbols \((x_1, x_2, x_3)\) are input to a communication channel, each with a probability of \(\frac{1}{3}\). The symbols \(y_1, y_2, \text{ and } y_3\) are received. If a symbol \(y_j\) is received then an error is made if a symbol \(x_i, i \neq j\), has been transmitted. The transition probability \(p(y_j|x_i) = q_{ij}\) are given by the following transition matrix:

\[
Q = \begin{bmatrix}
0.5 & 0.3 & 0.2 \\
0.4 & 0.3 & 0.3 \\
0.1 & 0.9 & 0
\end{bmatrix}
\]

(a) How much information \(H(Y)\) does one receive per symbol at receiving end?
(b) How much is the noise influence \(H(Y|X)\)?
(c) Calculate the error probability \(p(e|y_2)\) in the case when a symbol \(y_2\) is received.
(d) Calculate the average error probability from the point of view of the receiver, i.e., \(\sum_i p(e|y_i)p(y_i)\).
(e) Calculate the average error probability from the point of view of the sender, i.e., \(\sum_i p(e|x_i)p(x_i)\).
(f) Show that, in general, the average error probabilities are the same from the viewpoints of sender and receiver.
(g) Is the Fano inequality satisfied?