PROBLEM 1. Show that a cascade of \( n \) identical binary symmetric channels,
\[
X_0 \xrightarrow{\text{BSC }#1} X_1 \rightarrow \cdots \rightarrow X_{n-1} \xrightarrow{\text{BSC }#n} X_n
\]
each with raw error probability \( p \), is equivalent to a single BSC with error probability
\[
\frac{1}{2}(1 - (1 - 2p)^n)
\]
and hence that \( \lim_{n \to \infty} I(X_0; X_n) = 0 \) if \( p \neq 0, 1 \). Thus, if no processing is
allowed at the intermediate terminals, the capacity of the cascade tends to zero.

PROBLEM 2. Consider a memoryless channel with transition probability matrix \( P_{Y|X}(y|x) \),
with \( x \in \mathcal{X} \) and \( y \in \mathcal{Y} \). For a distribution \( Q \) over \( \mathcal{X} \), let \( I(Q) \)
denote the mutual information between the input and the output of the channel when the input distribution is \( Q \). Show
that for any two distributions \( Q \) and \( Q' \) over \( \mathcal{X} \),
\[
\text{(a)} \quad I(Q') \leq \sum_{x \in \mathcal{X}} Q'(x) \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x')Q'(x')} \right)
\]
\[
\text{(b)} \quad C \leq \max_x \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x')Q(x')} \right)
\]
where \( C \) is the capacity of the channel. Notice that this upper bound to the capacity
is independent of the maximizing distribution.

PROBLEM 3.

(a) Show that \( I(U;V) \geq I(U;V|T) \) if \( T, U, V \) form a Markov chain, i.e., conditional on
\( U \), the random variables \( T \) and \( V \) are independent.

Fix a conditional probability distribution \( p(y|x) \), and suppose \( p_1(x) \) and \( p_2(x) \) are two
probability distributions on \( \mathcal{X} \).

For \( k \in \{1, 2\} \), let \( I_k \) denote the mutual information between \( X \) and \( Y \) when the
distribution of \( X \) is \( p_k(\cdot) \).

For \( 0 \leq \lambda \leq 1 \), let \( W \) be a random variable, taking values in \( \{1, 2\} \), with
\[
\Pr(W = 1) = \lambda, \quad \Pr(W = 2) = 1 - \lambda.
\]

Define
\[
p_{W,X,Y}(w,x,y) = \begin{cases} 
\lambda p_1(x)p(y|x) & \text{if } w = 1 \\
(1 - \lambda)p_2(x)p(y|x) & \text{if } w = 2.
\end{cases}
\]

(b) Express \( I(X;Y|W) \) in terms of \( I_1, I_2 \) and \( \lambda \).

(c) Express \( p(x) \) in terms of \( p_1(x), p_2(x) \) and \( \lambda \).
(d) Using (a), (b) and (c) show that, for every fixed conditional distribution \( p_{Y|X} \), the mutual information \( I(X;Y) \) is a concave \( \cap \) function of \( p_X \).

**Problem 4.** Suppose \( Z \) is uniformly distributed on \([-1,1]\), and \( X \) is a random variable, independent of \( Z \), constrained to take values in \([-1,1]\). What distribution for \( X \) maximizes the entropy of \( X + Z \)? What distribution of \( X \) maximizes the entropy of \( XZ \)?

**Problem 5.** Random variables \( X \) and \( Y \) are correlated Gaussian variables:

\[
\begin{pmatrix}
X \\
Y
\end{pmatrix}
\sim \mathcal{N}_2\left(\begin{pmatrix}
0 \\
0
\end{pmatrix}; K = \begin{bmatrix}
\sigma_x^2 & \rho \sigma_x \sigma_y \\
\rho \sigma_x \sigma_y & \sigma_y^2
\end{bmatrix}\right).
\]

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

**Problem 6.** Suppose \( X \) and \( Y \) are independent geometric random variables. That is, \( p_X(k) = (1-p)^{k-1}p \) and \( p_Y(k) = (1-q)^{k-1}q \), \( \forall k \in \{1,2,\ldots\} \).

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

(b) Find \( H(2X + Y, X - 2Y) \)

Now consider two independent exponential random variables \( X \) and \( Y \). That is, \( p_X(t) = \lambda_X e^{-\lambda_X t} \) and \( p_Y(t) = \lambda_Y e^{-\lambda_Y t} \), \( \forall t \in [0,\infty) \).

(c) Find \( h(X,Y) \).

(d) Find \( h(2X + Y, X - 2Y) \)