This assignment consists of two pages.

1. The usual memoryless binary symmetric channel may be viewed as a channel of the form

\[ Y_n = X_n \oplus Z_n, \quad n = 1, 2, 3, \ldots, \]

where \( X_n \in \{0, 1\} \) for all \( n \), \((Z_n)\) is an i.i.d. Ber\((p)\) sequence independent of \((X_n)\), and \( \oplus \) denotes modulo-2 addition. We know that the capacity of this channel is \( 1 - H(p) \).

Now, consider a binary symmetric channel with memory, described as follows. At the beginning of time, a noise bit \( Z \) is randomly chosen according to a Ber\((p)\) distribution, and thereafter fixed for all time:

\[ Y_n = X_n \oplus Z, \quad n = 1, 2, 3, \ldots. \]

Determine the capacity of this channel, where we now define capacity to be

\[ C = \lim_{n \to \infty} \frac{1}{n} C_n, \]

with \( C_n = \max_{p(x^n)} I(X^n; Y^n) \), the maximum being taken over all pmfs \( p(x^n) \) on the input alphabet \( \{0, 1\}^n \).

2. Consider the function

\[ f(x) = \begin{cases} 
\frac{1}{x(\ln x)^2} & \text{if } x > e \\
0 & \text{otherwise}
\end{cases} \]

(a) Verify that \( f \) is a density function.

(b) Determine \( h(f) \).

3. (a) Let \( S \subseteq \mathbb{R} \), functions \( r_i(x) \), \( i = 1, 2, \ldots, m \), and constants \( \alpha_i \), \( i = 1, 2, \ldots, m \), all be given. Let \( \mathcal{F} \) represent the family of all densities \( f \) with the following properties:

- \( f(x) = 0 \) for all \( x \notin S \), i.e., \( \text{Supp}(f) \subseteq S \)
- \( \int_S f(x) \, dx = 1 \)
- \( \int_S r_i(x) f(x) \, dx = \alpha_i \), for \( i = 1, 2, \ldots, m \).

Now, let \( f^*(x) = \exp(\lambda_0 + \sum_{i=1}^{m} \lambda_i r_i(x)) \), where \( \lambda_0, \lambda_1, \ldots, \lambda_m \) are chosen so that \( f^* \) belongs to the family \( \mathcal{F} \). (Assume that such \( \lambda_i \)’s can indeed be chosen.) Prove that \( f^* \) uniquely maximizes the differential entropy \( h(f) \) over all densities \( f \in \mathcal{F} \).

(b) Let \( X \) be a non-negative random variable with \( \mathbb{E}[X] = \mu \), where \( \mu > 0 \) is a fixed constant. Show that \( h(X) \leq h(Z) \), where \( Z \sim \text{EXP}\left(\frac{1}{\mu}\right) \), i.e., \( Z \) has density \( f(z) = \frac{1}{\mu} e^{-z/\mu} \), \( z \geq 0 \).

(c) What happens if we remove the non-negative requirement on \( X \) in part (b)? In other words, what is \( \sup h(X) \), where the supremum is taken over all continuous random variables \( X \) with fixed mean \( \mathbb{E}[X] = \mu \).

4. Problem 8.8, Cover & Thomas, 2nd ed.

5. Problem 9.3, Cover & Thomas, 2nd ed.
7. Let $Y = X + Z$, where $X \sim \mathcal{N}(0, P)$, and $Z$ is a random variable independent of $X$, with mean 0 and variance $\sigma^2$. Given an observation of $Y$, we want to estimate $X$.

(a) Among linear estimators $\hat{X}(Y)$ of the form $aY + b$, determine the estimator that minimizes the mean squared error $\mathbb{E}[(X - \hat{X}(Y))^2]$. What is the resulting minimum mean squared error?

(b) Give an upper bound on $h(X|Y)$ using the result of part (a). When is this bound tight?

8. Consider any continuous channel with additive noise as follows:

$$Y_n = X_n + Z_n,$$

where the noise sequence $(Z_n)$ is iid with mean 0 and variance $\sigma^2$, and $(Z_n)$ is independent of $(X_n)$. Prove that the capacity, $C(P)$, of this channel, under an average input power constraint $P$, is greater than or equal to

$$\frac{1}{2} \log(1 + P/\sigma^2),$$

with equality iff the additive noise is iid $\mathcal{N}(0, \sigma^2)$.

Thus, among additive noise channels of fixed noise variance, the Gaussian channel has the least capacity. [Hint: Use the result of Problem 7(b).]