1. Entropy (7 points)

   a) The Rényi 2-entropy of a discrete random variable X with probability distribution P_X on $\mathcal{X}$ is $H_2(X) = -\log \sum_{x \in \mathcal{X}} P_X(x)^2$. Prove that this is a lower bound for the Shannon entropy.

   b) Characterize the set of probability distributions for which equality holds in this inequality. Justify your claim.

   c) Provide the definition of the conditional mutual information between the discrete random variables X and Z given Y.

   d) Provide an interpretation of $I(X : Z|Y) = 0$.

2. Entropy rates and data compression (6 points)

   a) When is a stochastic process called "stationary"?

   b) These two quantities are called the entropy rates of the stochastic process $X = (X_1, X_2, \ldots)$:

   $$H(X) = \lim_{n \to \infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n) \text{ and } H'(X) = \lim_{n \to \infty} H(X_n | X_1, X_2, \ldots, X_{n-1})$$

   Prove that for any stationary stochastic process these limits exist and that they are equal.

   c) Given a stationary stochastic process, provide a general lower bound on the average code word length of uniquely decodable codes.

   d) In which sense is there an inequality in the other direction? Provide a precise statement.

3. Shannon's noisy channel coding theorem (6 points)

   a) Compute the capacity of the channel given by the stochastic matrix $S = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{pmatrix}$.

   b) Show that $(\frac{2}{5}, \frac{3}{5})$ is the optimal input distribution for the channel given by the stochastic matrix

   $$S = \begin{pmatrix} \frac{p}{2} & p \\ \frac{1}{2} & 0 \\ \frac{1-p}{2} & 1-p \end{pmatrix}$$

   for any value of p.

4. Coding theory (6 points)

   a) Show that there exists a code over $GF(3)$ of length 11 and distance 3 with 3^6 codewords.

   b) Consider the set of univariate polynomials over $GF(2^4)$ of degree at most 3, and construct codewords by evaluating all such polynomials at the nonzero elements of $GF(2^4)$. How many errors can be corrected by this code?

   c) Suppose that we choose a basis of $GF(2^4)$ as a vector space over $GF(2)$, and represent the symbols of the codewords as those of a binary code by expanding in this basis. Can this code correct an error burst of length 21?