Problem 1. Let X_1 and X_2 be discrete random variables drawn according to probability mass functions $p_1(.)$ and $p_2(.)$ over the respective alphabets $\mathcal{X}_1 = \{1, 2, \ldots, m\}$ and $\mathcal{X}_2 = \{m+1, m+2, \ldots, n\}$. Let

$$X = \begin{cases} X_1 & \text{with probability } \alpha, \\ X_2 & \text{with probability } 1-\alpha, \end{cases}$$

a) Find $H(X)$ in terms of $H(X_1)$ and $H(X_2)$ and α.

b) Using the previous part, show that $2^{H(X)} \leq 2^{H(X_1)} + 2^{H(X_2)}$.

Problem 2. A discrete memoryless source emits a sequence of statistically independent binary digits with probabilities $p(1) = 0.015$ and $p(0) = 0.985$. The digits are taken 100 at a time and a binary codeword is provided for every sequence of 100 digits containing three or fewer ones.

a) Assuming that all codewords are the same length, find the minimum length required to provide codewords for all sequences with three or fewer ones.

b) Calculate the probability that a 100-bit sequence has three or fewer ones

c) Use Chebyshev's inequality to bound the probability of observing a source sequence for which no codeword has been assigned. Compare this bound with the actual probability computed in part (b).

Problem 3. Suppose that X, Y, Z are three random variables in such a way that $X = Y + Z$, where $X, Y \in \{0, 1, \ldots, m-1\}$.

a) Compare $H(X|Y)$ and $H(Z)$.

b) When is $H(X|Y)$ equal to $H(Z)$?

c) Let $U \leftrightarrow V \leftrightarrow (W, T)$ form a Markov chain in which W is some random variable. Prove that $I(U; W) - I(W; T) \leq I(U; V) - I(U; T)$.

Problem 4. A seven-game series between Nadal and Federer terminates as soon as either of the players wins four games. Let X be the random variable that represents the outcome of a World Series between Nadal (N) and Federer (F). Possible values of X are NNNN, NFFFNNF, and NNNFFFN. Let Y be the number of games played, which ranges from 4 to 7. Assuming that both players are equally matched and that the games are independent, calculate $H(X)$, $H(Y)$, $H(Y|X)$, and $H(X|Y)$.

Problem 5. Let $\{X_i\}$ be iid according to the distribution $p(x)$ and x takes the values from the set $\mathcal{X} = \{1, 2, \ldots, m\}$. Let $B^n = \{x^n \in \mathcal{X}^n : |\frac{1}{n}\sum_{i=1}^n x_i - E[X]| \leq \epsilon\}$. Prove that if n is large enough then:

$$H(X) - \epsilon - \frac{1}{n} \leq \frac{1}{n} \log |B^n \cap A_{(\epsilon)}^n| \leq H(X) + \epsilon.$$