Problem 1. Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution $p(x)$ taking values in a finite set $\mathcal{X}$. Thus, $p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i)$.

(a) Use the law of large numbers to show that

$$-\frac{1}{n} \log p(X_1, \ldots, X_n) \to H(X)$$

in probability.

Let $q(x_1, \ldots, x_n) = \prod_{i=1}^{n} q(x_i)$, where $q(x)$ is another probability distribution on $\mathcal{X}$.

(b) Evaluate

$$\lim_{n \to \infty} -\frac{1}{n} \log q(X_1, \ldots, X_n).$$

(c) Now evaluate the limit of the log-likelihood-ratio

$$-\frac{1}{n} \log \frac{q(X_1, \ldots, X_n)}{p(X_1, \ldots, X_n)}.$$

Problem 2. Assume $\{X_n\}_{n=1}^{\infty}$ and $\{Y_n\}_{n=1}^{\infty}$ are two i.i.d. processes (individually) with the same alphabet, with the same entropy rate $H(X_0) = H(Y_0) = 1$ and independent from each other. We construct two processes $Z$ and $W$ as follows:

- To construct the process $Z$, we flip a fair coin and depending on the result $\Theta \in \{0, 1\}$ we select one of the processes. In other words, $Z_n = \Theta X_n + (1 - \Theta) Y_n$.

- To construct the process $W$, we do the coin flip at every time $n$. In other words, at every time $n$ we flip a coin and depending on the result $\Theta_n \in \{0, 1\}$ we select $X_n$ or $Y_n$ as follows $W_n = \Theta_n X_n + (1 - \Theta_n) Y_n$.

(a) Are $Z$ and $W$ stationary processes? Are they i.i.d. processes?

(b) Find the entropy rate of $Z$ and $W$. How do they compare? When are they equal?

Recall that the entropy rate of the process $U$ (if exists) is $\lim_{n \to \infty} \frac{1}{n} H(U_1, \ldots, U_n)$.

Problem 3. Decode the string 10010011 that was encoded using the Lempel–Ziv algorithm with alphabet set $\mathcal{U} = \{a, l\}$.

Problem 4. Let the alphabet be $\mathcal{X} = \{a, b\}$. Consider the infinite sequence $X_1^{\infty} = ababababababab\ldots$

(a) What is the compressibility of $\rho(X_1^{\infty})$ using finite-state machines (FSM) as defined in class? Justify your answer.

(b) Design a specific FSM, call it $M$, with at most 4 states and as low a $\rho_M(X_1^{\infty})$ as possible. What compressibility do you get?
(c) Using only the result in point (a) but no specific calculations, what is the compressibility of $X_1^\infty$ under the Lempel–Ziv algorithm, i.e., what is $\rho_{LZ}(X_1^\infty)$?

(d) Re-derive your result from point (c) but this time by means of an explicit computation.

**PROBLEM 5.** We have shown in class that

$$\binom{n}{k} \leq 2^{nh_2(\frac{k}{n})}.$$ 

(a) Given $n \in \mathbb{N}_+$ and $n_1, n_2, \ldots, n_K \in \mathbb{N}$ such that $\sum_{i=1}^{n} n_i = n$, we define the quantity

$$\binom{n}{n_1, n_2, \ldots, n_K} = \frac{n!}{n_1!n_2! \cdots n_K!}.$$

Show that

$$\binom{n}{n_1, n_2, \ldots, n_K} \leq 2^{nh_2(p_1, p_2, \ldots, p_K)},$$

where $p_i = \frac{n_i}{n}$ and $h(p_1, \ldots, p_K) = -\sum_{i=1}^{K} p_i \log(p_i)$.

Let $U_1, U_2, \ldots$ be the letters generated by a memoryless source with alphabet $\mathcal{U} = \{u_1, u_2, \ldots, u_K\}$, i.e., $U_1, U_2, \ldots$ are i.i.d. random variables taking values in the alphabet $\mathcal{U}$ according to the distribution $q = \{q_1, q_2, \ldots, q_K\}$.

(b) We want to compress this source without any idea about its distribution. Describe an optimal universal code that achieves this goal. Give a proof of its optimality.

Hint: Use the same idea as for the binary source case.

(c) What if the source is not i.i.d. Will your code still be optimal?