Problem 1. For a stationary process $X_1, X_2, \ldots$, show that

(a) $\frac{1}{n}H(X_1, \ldots, X_n) \leq \frac{1}{n-1}H(X_1, \ldots, X_{n-1})$.

(b) $\frac{1}{n}H(X_1, \ldots, X_n) \geq H(X_n|X_{n-1}, \ldots, X_1)$.

Problem 2. Let $\{X_i\}_{i=-\infty}^{\infty}$ be a stationary stochastic process. Prove that

$$H(X_0|X_{-1}, \ldots, X_{-n}) = H(X_0|X_1, \ldots, X_n).$$

That is: the conditional entropy of the present given the past is equal to the conditional entropy of the present given the future.

Problem 3. A discrete memoryless source has alphabet 1, 2, where symbol 1 has duration 1 and symbol 2 has duration 2. The probabilities of 1 and 2 are $p_1$ and $p_2$ respectively. Find the value of $p_1$ that maximizes the source entropy per unit time, $H(X)/E[l_X]$, where $l_X$ is the duration of the symbol $x$. What is the maximum value of the entropy per unit time?

Problem 4. Define the type $P_x$ (or empirical probability distribution) of a sequence $x_1, \ldots, x_n$ be the relative proportion of occurrences of each symbol of $X$; i.e.,

$$P_x(a) = \frac{N(a|x)}{n}$$

for all $a \in X$, where $N(a|x)$ is the number of times the symbol $a$ occurs in the sequence $x \in X^n$.

(a) Show that if $X_1, \ldots, X_n$ are drawn i.i.d. according to $Q(x)$, the probability of $x$ depends only on its type and is given by

$$Q^n(x) = 2^{-n(H(P_x) + D(P_x||Q))}.$$

Hint: Start by showing the following:

$$Q^n(x) = \prod_{i=1}^{n} Q(x_i) = \prod_{a \in X} Q(a)^{N(a|x)} = \prod_{a \in X} Q(a)^{nP_x(a)}$$

Define the type class $T(P)$ as the set of sequences of length $n$ and type $P$:

$$T(P) = \{x \in X^n : P_x = P\}.$$

For example, if we consider binary alphabet, the type is defined by the number of 1’s in the sequence and the size of the type class is therefore $\binom{n}{k}$.

(b) Show for a binary alphabet that

$$|T(P)| = 2^{nH(P)}.$$  \hspace{1cm} (1)

We say that $a_n \approx b_n$, if $\lim_{n \to \infty} \frac{1}{n} \log \frac{a_n}{b_n} = 0$. 


Hint: Prove that
\[
\frac{1}{n+1}2^{nh_2(\frac{k}{n})} \leq \binom{n}{k} \leq 2^{nh_2(\frac{k}{n})}.
\]

\(h_2(\cdot)\) denotes the binary entropy function. To derive the upper bound, start by proving
\[
1 \geq \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} = \binom{n}{k} 2^n \left(\frac{k}{n} \log \frac{k}{n} + \frac{n-k}{n} \log \frac{n-k}{n}\right).
\]

To derive the lower bound, start by proving
\[
1 = \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \leq (n+1) \max_j \binom{n}{j} p^j (1-p)^j,
\]
then take \(p = k/n\) and show that the maximum occurs for \(j = k\).

(c) Use (a) and (b) to show that
\[
Q^n(T(P)) \geq 2^{-nD(P||Q)}.
\]

Note: \(D(P||Q)\) is the informational divergence (or Kullback-Leibler divergence) between two probability distributions \(P\) and \(Q\) on a common alphabet \(\mathcal{X}\) and is defined as
\[
D(P||Q) = \sum_{a \in \mathcal{X}} P(a) \log \frac{P(a)}{Q(a)}.
\]

Recall that we have already seen the non-negativity of this quantity in the class.

Problem 5. Suppose we bet our fortune at a casino game which multiplies our fortune by a random variable \(X\), with
\[
\Pr(X = 1/4) = \Pr(X = 2) = 1/2.
\]

We play the game repeatedly, starting with an initial fortune \(F_0 = 1\), betting our entire fortune at each time. The value of \(X\) at the \(i\)th game is denoted by \(X_i\). These values are independent and distributed as \(X\).

(a) What is the expected value \(f_n = E[F_n]\), our fortune after \(n\) plays? How does \(f_n\) behave as \(n\) gets large?

(b) What is \(l_n = E[\log_2 F_n]\)? How does \(2^{l_n}\) behave as \(n\) gets large?

(c) Does \(F_n\) concentrate around \(f_n\) or \(2^{l_n}\)? [Hint: does the law of large numbers apply to \(F_n\) or to \(\log_2 F_n\)?]

(d) With this ‘bet all we have’ strategy do we get rich or poor?

(e) If we had kept a fraction \(r\) of our fortune in reserve at each play, could we have done better? What is the best value of \(r\) to maximize \(\lim_{n \to \infty} \frac{1}{n} \log_2 F_n\), the ‘rate of growth’ of fortune.