Homework Set #1
Entropy rate of stationary processes,
directed information and causal conditioning

1. *Monotonicity of entropy per element.*

   For a stationary stochastic process $X_1, X_2, \ldots, X_n$, show that

   (a)
   $$\frac{H(X_1, X_2, \ldots, X_n)}{n} \leq \frac{H(X_1, X_2, \ldots, X_{n-1})}{n-1}.$$

   (b)
   $$\frac{H(X_1, X_2, \ldots, X_n)}{n} \geq H(X_n|X_{n-1}, \ldots, X_1).$$

2. *Pairwise independence.* Let $X_1, X_2, \ldots, X_{n-1}$ be i.i.d. random variables taking values in $\{0, 1\}$, with $\Pr\{X_i = 1\} = \frac{1}{2}$. Let $X_n = 1$ if $\sum_{i=1}^{n-1} X_i$ is odd and $X_n = 0$ otherwise. Let $n \geq 3$.

   (a) Show that X_i and X_j are independent, for $i \neq j, i, j \in \{1, 2, \ldots, n\}$.
   (b) Find $H(X_i, X_j)$, for $i \neq j$.
   (c) Find $H(X_1, X_2, \ldots, X_n)$. Is this equal to $nH(X_1)$?

3. *Cesáro mean.* Prove that if $\lim a_n = a$ and $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$, then $\lim b_n = a$.

4. *Stationary processes.* Let $\ldots, X_{-1}, X_0, X_1, \ldots$ be a stationary (not necessarily Markov) stochastic process. Which of the following statements are true? State true or false. Then either prove or provide a counterexample. Warning: At least one answer is false.

   (a) $H(X_n|X_0) = H(X_{-n}|X_0)$.
   (b) $H(X_n|X_0) \geq H(X_{n-1}|X_0)$.
   (c) $H(X_n|X_1^{n-1}, X_{n+1})$ is nonincreasing in n.

5. *Directed Information and causal conditioning* Directed information is denoted as $I(X^n \to Y^n)$ and is defined as

$$I(X^n \to Y^n) \triangleq \sum_{i=1}^{n} I(X^i; Y_i | Y^{i-1}).$$

Causal conditioning is denoted as $p(y^n || x^{n-d})$ and is defined as

$$p(y^n || x^{n-d}) \triangleq \prod_{i=1}^{n} p(y_i | y^{i-1}, x^{i-d})$$

a. Prove that

i.
$$I(X^n \to Y^n) \triangleq E\left[\log \frac{p(Y^n || X^n)}{p(Y^n)}\right]$$

ii.
$$p(y^n, x^n) = p(y^n || x^n) p(x^n || y^{n-1}),$$

iii.
$$I(X^n \to Y^n) \geq 0$$

and equals zero if and only if $p(y^n || x^n) = p(y^n)$

iV.
$$I(X^n; Y^n) = I(X^n \to Y^n) + I(0Y^{n-1} \to X^n),$$

where the term , $0Y^{n-1}$, denotes the concatenation of 0 (Null) to the sequence Y^{N-1}, i.e. $(0, Y_1, Y_2, ...Y_{n-1})$.

b. Prove that in general

$$I(X^n \to Y^n) \leq I(X^n; Y^n)$$

and equality holds if and only if $p(x^n || y^{n-1}) = p(x^n)$.

c. Suggest (without proof) properties of directed information and causal conditioning. Notice that every property that holds for directed information should also hold for mutual information. A student who suggests a nontrivial property for directed information or causal conditioning that was not mention by any other student and not mentioned in the literature will receive 20 points bonus for the HW assignment.