Information Theory
Exercises

1. A homogenous Markov chain has three states: A, B, and C. From state A it goes to state B with probability 1. From state B it goes to state C with probability 1/3 and stays in state B with probability 2/3. From state C it goes to state A with probability 2/3 and stays in state C with probability 1/3. Determine the entropy of the source formed by this Markov chain. (That is, the source emits a symbol after each state transition of the given Markov chain and the output is simply the new state.)

   Sketch of Solution: Let the stationary distribution be (a, b, c). Then we can write $a = \frac{2}{3}c, b = a + \frac{2}{3}b, c = \frac{1}{3}b + \frac{1}{3}c$, plus $a + b + c = 1$ (Note that this system of equations is redundant.) Solving it, we obtain: $a = \frac{2}{11}, b = \frac{6}{11}, c = \frac{3}{11}$. The entropy of the source is then

$$H(X) = \frac{2}{11} \cdot 0 + \frac{6}{11} \cdot h(1/3) + \frac{3}{11} \cdot h(1/3) = \frac{9}{11}(\log 3 - 2/3).$$

2. Let $\mathbf{X} = (X_1, X_2, \ldots)$ be a stationary source with entropy $H(\mathbf{X})$. Decide whether the entropy of the following sources exists and determine it if it does.

   a) $\mathbf{X}_a = (X_1, X_1, X_2, X_2, X_3, X_3, \ldots)$ (all random variables are repeated once)

   b) $\mathbf{X}_b = (X_1, X_1, X_2, X_3, X_3, X_4, X_5, X_5, X_6, \ldots)$ (only the odd indexed random variables are repeated)

   c) $\mathbf{X}_c = (X_1, X_2, X_2, X_3, X_3, X_3, X_4, X_4, X_4, X_4, \ldots)$ (the random variable with index i is repeated i times)

   Sketch of Solution:

   a)

$$\begin{aligned} H(X_a) &= \lim_{n\to\infty} \frac{1}{2n} H(X_1, X_1, X_2, X_2, \ldots, X_n, X_n) = \\ &\frac{1}{2} \lim_{n\to\infty} \frac{1}{n} H(X_1, X_1, X_2, X_2, \ldots, X_n, X_n) = \\ &\frac{1}{2} \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n) = \frac{1}{2} H(X). \end{aligned}$$

   b)

$$\begin{aligned} H(X_b) &= \lim_{n\to\infty} \frac{1}{1.5n} H(X_1, X_1, X_2, X_3, X_3, X_4, X_4 \ldots, X_n) = \\ &\frac{2}{3} \lim_{n\to\infty} \frac{1}{n} H(X_1, X_1, X_2, X_3, X_3 \ldots, X_n) = \\ &\frac{2}{3} \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n) = \frac{2}{3} H(X). \end{aligned}$$

   c)

$$\begin{aligned} H(X_b) &= \lim_{n\to\infty} \frac{2}{n(n+1)} H(X_1, X_2, X_2, X_3, X_3, X_3 \ldots, X_n, \ldots, X_n) = \\ &\lim_{n\to\infty} \frac{2}{n+1} \frac{1}{n} H(X_1, X_2, X_2, X_3, X_3, X_3 \ldots, X_n, \ldots, X_n) = \\ &\lim_{n\to\infty} \frac{2}{n+1} \frac{1}{n} H(X_1, X_2, \ldots, X_n) = \\ &\lim_{n\to\infty} \frac{2}{n+1} H(X) = 0. \end{aligned}$$

3. Let the random variable X have density function $f(x)$ given as follows.

$$f(x) = x + 1 \text{ if } x \in [-1, 0], \quad f(x) = -x + 1 \text{ if } x \in [0, 1],$$

and $f(x)$ is 0 outside the interval $[-1, 1]$. Perform the first iteration of the Lloyd-Max algorithm for a two-level quantizer of the variable X starting with initial quantization values $-0.5, 0.5$.

Solution: The arithmetic mean of the two quantization levels is 0, so we have $B_1 = [-1, 0), B_2 = [0, 1]$.

$$\frac{\int_{-1}^{0} xf(x)dx}{\int_{-1}^{0} f(x)dx} = \frac{\int_{-1}^{0} x^2 + xdx}{1/2} = 2\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_{-1}^{0} = -\frac{1}{3}.$$

Similarly,

$$\frac{\int_{0}^{1} xf(x)dx}{\int_{0}^{1} f(x)dx} = \frac{\int_{0}^{1} -x^2 + 1dx}{1/2} = 2\left[-\frac{x^3}{3} + \frac{x^2}{2}\right]_{0}^{1} = \frac{1}{3}.$$

Remark: The B_i's do not change, so this is a Lloyd-Max quantizer.

4. A source $\mathbf{X} = X_1, X_2, \ldots$ works as follows. Each X_i is equal to either 0 or 1, $Prob(X_1 = 0) = Prob(X_1 = 1) = 1/2$ and similarly $Prob(X_2 = 0) = Prob(X_2 = 1) = 1/2$. For $i \geq 3$ the rule is the following. If $X_{i-1} = X_{i-2}$ then $X_i = 1 - X_{i-1}$ for sure (that is, with probability 1). If $X_{i-1} \neq X_{i-2}$, then $X_i = 0$ and $X_i = 1$ has equal probability, that is $Prob(X_i = 0|X_{i-1} \neq X_{i-2}) = Prob(X_i = 1|X_{i-1} \neq X_{i-2}) = \frac{1}{2}$. Give the entropy of the source $\mathbf{X}$ if it exists.

Sketch of Solution: We know by a theorem that

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

In this case

$$\lim_{n\to\infty} H(X_n|X_1, X_2, \ldots, X_{n-1}) = \lim_{n\to\infty} H(X_n|X_{n-1}, X_{n-2}),$$

since the equality or non-equality of X_{n-1} and X_{n-2} determines the distribution for X_n. We can think about the system as a Markov chain with two states A and B, where A means that the last two outputs were equal, B means that they were not. Then the transition probabilities are: from A the system goes to B with probability 1, from B it goes to both A and B with probability 1/2. Calculating the stationary distribution the usual way we get that $Prob(A) = 1/3, Prob(B) = 2/3$. Thus

$$\lim_{n\to\infty} H(X_n|X_{n-1}, X_{n-2}) =$$

$$\lim_{n\to\infty} (Prob(X_{n-1} = X_{n-2})H(X_n|X_{n-1} = X_{n-2}) +$$

$$+ Prob(X_{n-1} \neq X_{n-2})H(X_n|X_{n-1} \neq X_{n-2})) =$$

$$1/3 \cdot 0 + 2/3 \cdot h(1/2) = 2/3.$$

5. We have two channels, both with input alphabet $\{0, 1, 2\}$ and output alphabet $\{0, 1\}$.

The probabilities describing channel A are as follows:

$$W_A(0|0) = 1, W_A(1|0) = 0$$

$$W_A(0|1) = 0, W_A(1|1) = 1$$

$$W_A(0|2) = 1/2, W_A(1|2) = 1/2$$

The probabilities describing channel B are as follows:

$$W_B(0|0) = 1, W_B(1|0) = 0$$

$$W_B(0|1) = 1, W_B(1|1) = 0$$

$$W_B(0|2) = 0, W_B(1|2) = 1$$

a) Give the capacity of both channels

b) Give the capacity of the channel we obtain by using the above two channels together. This means we send every bit via both, and the received two bits (one on each channel) thus provide four possible outputs: $00, 01, 10, 11$

Sketch of Solution: a) Both channels have capacity $\log 2 = 1$. Neither can be larger as $C = \max I(X, Y) = \max(H(Y) - H(Y|X)) \leq \log 2$, since $H(Y) \leq \log 2$ (as the output is binary) and $H(Y|X) \geq 0$.

This maximum value is attainable in both cases. For channel A by choosing the input probability of 2 to be 0 and equal (that is 1/2) for 0 and 1. For channel B we attain it if the input probability is 1/2 for 2. (It will not matter how we distribute the remaining 1/2 probability on 0 and 1, but the calculation is simplest if we concentrate it on, say, 0.)

b) For the two channels used together a 0 input results in 00 at the output, input 1 in 01 for sure and input 2 in 10 or 11. So the input can be told from the output with certainty, thus $H(X|Y) = 0$. So $I(X, Y) = H(X) - H(X|Y)$ is largest when $H(X)$ is largest. This is when the input distribution is uniform on the three input letters. Thus the capacity of this combined channel is $\log 3$, that is larger than that of the individua ones.