Exercise sheet n° 1:

1-1. A discrete random variable X has a uniform probability distribution $p(x)$ over the set $\mathcal{X}$ of cardinality $|\mathcal{X}| = m$.

(a) Give an example for $\mathcal{X}$ and calculate $p(x)$ for $x \in \mathcal{X}$.

(b) Determine the entropy $H(X)$. If $m = 64$, what is $H(X)$?

(c) How many bits are needed to enumerate the alphabet $\mathcal{X}$ without encoding, where $m = 64$?

(d) How many symbols taken from a quaternary alphabet are necessary to enumerate the alphabet $\mathcal{X}$ without encoding, where $m = 64$?

(e) Show that $H(X)$ is greater than the entropy of any other random variable Y, where Y has values on the same set $\mathcal{X}$. *(Read the recall on Lagrange multipliers.)*

1-2. A random variable X has an alphabet $\mathcal{X} = \{A, B, C, D\}$.

(a) Calculate $H(X_P)$ associated to probabilities
$P(X) = \{1/2, 1/4, 1/8, 1/8\}$.

(b) Calculate $H(X_Q)$ associated to probabilities $Q(X) = \{1/4, 1/4, 1/4, 1/4\}$.

(c) We define a binary code:
{C($X = A$)=0, C($X = B$)=10, C($X = C$)=110, C($X = D$)=111}.
Calculate the expected length (in bits) of the codewords of X are distributed according to $P(X)$ and $Q(X)$

(d) Compare the four obtained results.

1-3. A coin (with same probability for "head" or "tail") is flipped repeatedly until one obtains "head". Let the random variable X be given by the number of flips until one obtains "head" for the first time.

(a) Calculate $H(X)$.

(b) One repeatedly flips the coin until "head" is obtained for the first time. Find a sequence of questions of type "yes/no" to determine the value of X. Compare the entropy of X with the expected length of the sequence of questions necessary to fully determine X.

Note:

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}, \qquad \sum_{n=1}^{\infty} n r^n = \frac{r}{(1-r)^2}, \qquad r < 1.$$

1-4. Assume X and Y to be random variables. What is the inequality relating $H(X)$ and $H(Y)$ if,

(a) $Y = 2^X$, with $\mathcal{X} = \{0, 1, 2, 3, 4\}$ associated to probabilities $\{1/2, 1/4, 1/8, 1/16, 1/16\}$.

(b) $Y = \cos(X)$, with $\mathcal{X} = \{0, \pi/2, \pi, 3\pi/2, 2\pi\}$ associated to probabilities $\{1/3, 1/3, 1/12, 1/12, 1/6\}$.

(c) Show that the entropy of any function $f(X)$ of the random variable X is less or equal than the entropy of X. In order to prove this, it is useful to calculate the join entropy $H(X, f(X))$. In which case the inequality $H(f(X)) \leq H(X)$ is saturated?

1-5. Let X and Y be two random variables, taking values $x_1, x_2, \cdots, x_r$ and $y_1, y_2, \cdots, y_s$. Furthermore, we define a random variable $Z = X + Y$.

(a) Show that $H(Z|Y) = H(X|Y)$ and $H(Z|X) = H(Y|X)$. Deduce that if X and Y are independently distributed, then $H(X) \leq H(Z)$ and $H(Y) \leq H(Z)$. Thus, summing two random variables can only increase the uncertainty.

(b) Give an example for X and Y (necessarily correlated) such that $H(X) > H(Z)$ and $H(Y) > H(Z)$.

(c) In which case the equality $H(Z) = H(X) + H(Y)$ is satisfied?

Optional: Give an example of a distribution of two random variables X and Y whose correlation coefficient is zero, although they are not independent. Show that in this case $H(X{:}Y) \neq 0$, showing that the mutual entropy is a better measure of the dependence of X and Y. Correlation coefficient: $r = \frac{<xy> - <x><y>}{\sqrt{<x^2> - <x>^2}\sqrt{<y^2> - <y>^2}}$

Recall:

1. Change of the base of logarithms

   $\log_a N = \log_b N / \log_b a$

2. Lagrange theorem:

   - We want to maximize
     $u(x_1, x_2, ..., x_n)$ with k constraints $g_j(x_1, x_2, ..., x_n) = a_j$, $j = 1, \cdots, k$.
   - We introduce the Lagrangian:
     $L(x_1, ..., x_n, \lambda_1, ..., \lambda_k) = u(x_1, x_2, ..., x_n) + \sum_{j=1}^{k} \lambda_j [a_j - g_j(x_1, x_2, ..., x_n)]$.
   - Condition on an extremum:
     $\forall i, \frac{\partial L}{\partial x_i} = 0$.
   - If u is concave, the solution is a maximum.

Web site:

http://quic.ulb.ac.be/teaching/