

INFORMATION AND CODING THEORY
Exercise Sheet 1

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

- (a) Give an example for \mathcal{X} and calculate $p(x)$ for $x \in \mathcal{X}$.
- (b) Determine the entropy $H(X)$ of X . If $m = 64$, what is $H(X)$?
- (c) How many bits are needed to enumerate the alphabet of \mathcal{X} without encoding, when $m = 64$?
- (d) How many symbols taken from a quaternary alphabet are necessary to enumerate \mathcal{X} without encoding, when $m = 64$?
- (e) Show that $H(X)$ is greater than the entropy of any other random variable Y , when Y takes values from the same set as \mathcal{X} . (Read the reminder on Lagrange multipliers.)

Exercise 2. A random variable X is defined over the set $\mathcal{X} = \{A, B, C, D\}$.

- (a) Calculate $H(X_p)$ with associated probability distribution $p(X) = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}$.
- (b) Calculate $H(X_q)$ with associated probability distribution $q(X) = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$.
- (c) Consider the binary code defined by $\{C(X=A) = 0, C(X=B) = 10, C(X=C) = 110, C(X=D) = 111\}$. Calculate the expected length (in number of bits) of the codewords of X , when X obeys the distribution $p(X)$ and $q(X)$ respectively.
- (d) Compare the four results.

Exercise 3. A fair coin 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) Find a sequence of “yes/no” questions 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 .

Hint:

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r},$$

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

Exercise 4. Let X and Y be two random variables. What is the inequality relating $H(X)$ and $H(Y)$ if,

- (a) $Y = 2^X$, with $\mathcal{X} = \{0, 1, 2, 3, 4\}$ with associated probability distribution $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}\}$.
- (b) $Y = \cos(X)$, with $\mathcal{X} = \{0, \pi/2, \pi, 3\pi/2, 2\pi\}$ with associated probability distribution $\{\frac{1}{3}, \frac{1}{3}, \frac{1}{12}, \frac{1}{12}, \frac{1}{6}\}$.
- (c) Show that the entropy of any function $f(X)$ of the random variable X is less than or equal to the entropy of X . In order to prove this, it is useful to calculate the joint entropy $H(X, f(X))$. In which case the inequality $H(f(X)) \leq H(X)$ is saturated?

Exercise 5. Let X and Y be two random variables, which take the values x_1, x_2, \dots, x_r and y_1, y_2, \dots, y_s . Furthermore, 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 the two random variables can only increase the uncertainty.
- (b) Give an example for X and Y (which should be correlated) such that $H(X) > H(Z)$ and $H(Y) > H(Z)$.
- (c) In which case is the equality $H(Z) = H(X) + H(Y)$ satisfied?

Exercise 6. (Optional) Give an example of a distribution of two random variables X and Y whose correlation coefficient r (as defined below) is zero, even though they are not independent. Show that in this case $H(X:Y) \neq 0$, which shows that the mutual entropy is a better measure of the dependence of X and Y .

Definition. Correlation coefficient

$$r = \frac{\langle x \cdot y \rangle - \langle x \rangle \cdot \langle y \rangle}{\sqrt{\langle x^2 \rangle - \langle x \rangle^2} \sqrt{\langle y^2 \rangle - \langle y \rangle^2}}$$

Reminders

- (a) How to change the base in logarithm:

$$\log_a b = \log_x b / \log_x a.$$

- (b) Lagrange theorem:

- To maximize $u(x_1, x_2, \dots, x_n)$ with k constraints $g_j(x_1, x_2, \dots, x_n) = a_j$, $j = 1, \dots, k$.
- Introduce the Lagrangian:
 $L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_k) = u(x_1, x_2, \dots, x_n) + \sum_{j=1}^k \lambda_j [a_j - g_j(x_1, x_2, \dots, x_n)]$.
- With the condition of an extremum being:
 $\forall i, \frac{\partial L}{\partial x_i} = 0$.
- If u is concave, the solution is a maximum.

The exercises and solutions are available at <http://quic.ulb.ac.be/teaching>