TUTORIAL I

1 Weighing problem

You are given 12 balls, all equal in weight except for one that is either heavier or lighter. You are also given a classical two-pan balance which allows you only to compare two subsets of balls (you are not given any reference weight). Your task is to design a strategy to determine which is the odd ball *and* whether it is heavier or lighter, using as few uses of the balance as possible.

1. What is the amount of uncertainty of a configuration?
2. How much information on average can a single use of the balance give? What is the minimum number of weighing one can hope to achieve?
3. Show that if we start by weighing balls 1-6 against 7-12, we don't get enough information to achieve the optimal solution.
4. Describe an optimal strategy.
5. Compute the exact information obtained during the process depending on the result of the second round (3 or 2 remaining situations).

2 A realistic find query

We consider a list of 32 elements and we want to test if a given element z belongs to the list or not. We assume that the probability that the element belongs to the list is $1/2$, and that all the positions within the list are equiprobable. Our strategy is to test the first element, then the second element, ... until the wanted element is found or the end of the list is reached. We denote by F the random variable which is equal to 1 if and only if z is in the list, 0 otherwise.

1. Compute the entropy of F.
2. We denote by L_1 the random variable corresponding to the result of the first test. Compute the entropy of L_1.
3. Compute the distribution of the joint variable (F, L_1), and give the joint entropy $H(F, L_1)$.
4. Compute the conditional entropy $H(F|L_1)$.
5. We denote by $L_1, \dots, L_n$ the result of the successive tests. Compute directly the conditional entropy $H(F|L_1, \dots, L_n)$.
6. If we plug $n = 16$ in the previous solution, we find $0.689 > \frac{1}{2}$. Is it reasonable ? What is the value of $H(F|L_1, \dots, L_{32})$?

3 Axiomatic approach to the Shannon entropy

If we require certain properties of our uncertainty measure, then it uniquely specifies the Shannon entropy. Let $\Delta_m = \{(p_1, \dots, p_m) \in \mathbb{R}^m : p_i \geq 0, \sum_i p_i = 1\}$ be the set of distributions on m elements. Let our uncertainty measure $H_m : \Delta_m \to \mathbb{R}$ be a sequence of functions satisfying the following desirable properties

1. Symmetry: For any $m \geq 1$ and any permutation π of $\{1, \dots, m\}$, $H_m(p_1, \dots, p_m) = H_m(p_{\pi(1)}, \dots, p_{\pi(m)})$
2. Normalization: $H_2(\frac{1}{2}, \frac{1}{2}) = 1$
3. Continuity: For any $m \geq 1$, H_m is a continuous function
4. Grouping: For any $m \geq 2$,
$$H_m(p_1, \dots, p_m) = H_{m-1}(p_1 + p_2, p_3, \dots, p_m) + (p_1 + p_2) H_2(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2})$$
5. Monotonicity: We have $H_m(\frac{1}{m}, \dots, \frac{1}{m}) \leq H_{m+1}(\frac{1}{m+1}, \dots, \frac{1}{m+1})$

Prove that $H_m(p_1, \dots, p_m) = -\sum_{i=1}^m p_i \log_2 p_i$.

You can proceed in the following way. Let $g(m) = H_m(\frac{1}{m}, \dots, \frac{1}{m})$.

1. Show that $g(n \cdot m) = g(n) + g(m)$.
2. Conclude that $g(m) = \log_2 m$. (Hint: for any n, let ℓ_n be such that $2^{\ell_n} \leq m^n \leq 2^{\ell_n + 1}$, show that $\frac{\ell_n}{n} \leq g(m) \leq \frac{\ell_n + 1}{n}$).
3. Use this to compute the value of $H_2(p, 1-p)$.
4. Conclude with H_m.

4 Rényi entropy

The Rényi entropy of order α, where $0 \leq \alpha < 1$, is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \Big(\sum_{i=1}^n p_i^\alpha \Big).$$

where X is a discrete random variable taking value in $\{1, 2, \dots, n\}$ each with probability $p_i = \Pr[X = i]$ for $i = 1 \dots n$. To define $\alpha = 0$ setting, we say that $0^0 = 0$.

1. Show that Rényi entropy is non-increasing function of α.
2. What is the value of H_0 and H_1 (here H_1 is defined as Rényi entropy when $\alpha \to 1$).
3. Show that H_α is concave function of the distribution $(p_1, \dots, p_n)$.