# Introduction à la théorie de l'information

## TD 4 : Information élémentaire et entropie d'une source

Université Paris 13, Institut Galilée
Master Ingénierie et Innovations en Images et Réseaux - 1ère année

2017-2018

## Exercice 1 : Quelques calculs d'entropies

1. On considère une urne contenant 4 boules rouges, 2 bleues, 1 verte et 1 jaune. Calculer son entropie.
2. On considère la phrase **INSTITUT_GALILEE** Calculer son entropie.

## Exercice 2 : Entropies d'un système de communication

On considère deux sources (modélisées par les variables aléatoires $X$ et $Y$) émettant simultanément des symboles. La loi conjointe des deux variables aléatoires $X$ et $Y$ est donnée par le tableau suivant:

| $X$ \ $Y$ | **1** | **2** | **3** | **4** |
|---|---|---|---|---|
| **1** | $\frac{1}{8}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{4}$ |
| **2** | $\frac{1}{16}$ | $\frac{1}{8}$ | ??? | 0 |
| **3** | $\frac{1}{32}$ | $\frac{1}{32}$ | $\frac{1}{16}$ | 0 |
| **4** | $\frac{1}{32}$ | $\frac{1}{32}$ | $\frac{1}{16}$ | 0 |

1. Il manque une valeur dans ce tableau : que vaut-elle ?
2. Déterminer la loi marginale $p_X(x)$ associée à $X$. Même question concernant $Y$.
3. Calculer les entropies $H(X)$ et $H(Y)$. Laquelle apporte le plus d'information ?

## Exercice 3 : Entropies relatives

1. On considère deux variables aléatoires $X$ et $X'$ à valeurs dans $\{1, 2, 3\}$. On notera $p(x) = P_X(X = x)$ et $q(x) = P_{X'}(X' = x)$et on donne le tableau suivant :

| $x$ | $p(x)$ | $q(x)$ |
|---|---|---|
| **1** | $\frac{1}{2}$ | $\frac{1}{3}$ |
| **2** | $\frac{1}{4}$ | $\frac{1}{3}$ |
| **3** | $\frac{1}{4}$ | $\frac{1}{3}$ |

   (a) Calculer $H(X)$ et $H(X')$. Commenter.

   (b) Calculer $D_{KL}(p||q)$ et $D_{KL}(q||p)$. Que constate-on ?

2. Donner un exemple de deux distributions de probabiltés $p(x)$ et $q(x)$ associées à une variable aléatoire binaire telle que $p(x) \neq q(x)$ et $D_{KL}(p||q) = D_{KL}(q||p)$

## Exercice 4 : Entropie et questions

On considère $n$ boîtes numérotées de 1 à $n$. On suppose qu'un trésor est caché dans l'une des boîtes et que le joueur n'a aucune information sur la boite contenant le trésor. Pour trouver la boîte, il a le droit de poser des questions auxquelles le présentateur ne répond par OUI ou par NON.

1. On suppose que $n = 8$. Combien de questions sont nécessaires pour que le joueur soit sur de retrouver la boite gagnante ?

2. Etablir le meme raisonnement pour 16 boites et 32 boites, et postuler d'un résultat pour $n$ boites.

3. On considère maintenant $n = 8$ boites réparties en 3 catégories contenant respectivement $n_1 = 4$, $n_2 = 2$ et $n_3 = 2$ boites :

   - Catégorie 1 : Boites 1, 2, 3, 4
   - Catégorie 2 : Boites 5, 6
   - Catégorie 3 : Boites 7, 8

   (a) En utilisant le postulat de la question précédente, donner pour chaque catégorie le nombre de questions à poser pour déterminer la boite gagnante.

   (b) Déterminer les probabilités $p_i$ que la boite gagnante soit dans la catégorie $i$

   (c) En déduire le nombre moyen de questions que l'on doit poser si on connait déjà la catégorie.

   (d) On considère une source d'information qui nous donne l'information de la catégorie. Montrer que l'information $I$ qu'elle fournit nous permet de passer de $\log_2(n)$ questions à $\displaystyle\sum_{i=1}^{3} p_i \log_2(n_i)$ questions.

   (e) En déduire une autre interprétation de l'entropie.