ESPCI Deuxième année Année 1999-2000
Méthodes Mathématiques
Cyrille Barreteau
cyrille.barreteau@cea.fr
http ://www-drecam.cea.fr/Phocea/Membres/Cours/index.php

Tutorat no 1
Entropie et théorie de l'information.

Résumé

Dans ce tutorat nous abordons une grandeur fondamentale qui joue un rôle essentiel dans de multiples domaines des mathématiques, de la physique et de la théorie de l'information : l'entropie. Nous commençons par en étudier les propriétés fondamentales pour ensuite l'appliquer à la théorie de l'information et à la physique.

1 Définition et propriétés fondamentales

1.1 Définition

On considère une situation Π dans laquelle N évènements e_n peuvent se produire avec une probabilité p_n. On définit la grandeur S appelée entropie (ou aussi information I) mesurant la quantité d'information par

$$S(\Pi) = S(p1, p2, \cdots, p_N) = -\lambda \sum_n p_n \log p_n$$

λ est une constante positive que l'on peut choisir à sa convenance. En mécanique statistique on prend $\lambda = k_B$ la constante de Boltzmann ($k_B = 1.3810^{-23} J/K$), par contre en théorie de l'information où l'unité est le "bit" la constante est sans dimension et vaut $\lambda = 1/\log 2$.

1.2 Les postulats fondamentaux

La définition de l'entropie peut se déduire de quelques postulats fondamentaux que nous allons énoncer.
Dans ce qui suit il vous est demandé de vérifier que la définition de l'entropie satisfait bien ces postulats.

a)**Positivité et symétrie** :

$$S \geq 0 \quad \text{et} \quad S(p_1, p_2, \cdots, p_i, \cdots, p_j, \cdots, p_N) = S(p_1, p_2, \cdots, p_j, \cdots, p_i, \cdots, p_N) \quad \forall i, j$$

b) **Minimum** : L'entropie atteint son minimum pour un cas pur (un des p_i égal à un).

c) **Maximum** : N étant fixé l'entropie atteint son maximum pour N évènements équiprobables.

d) **Croissance** : Dans le cas d'évènement équiprobables l'entropie croit avec le nombre de ces évènements.

f) **Additivité** :
i) On considère deux ensembles d'évènements indépendants $\Pi = (e_n, n = 1, \cdots, N)$ et $\Pi' = (e'_{n'}, n' = 1, \cdots, N')$ de probabilités $(p_n, n = 1, \cdots, N)$ et $(p'_{n'}, n' = 1, \cdots, N')$. La probabilité d'occurence de l'évènement composé $(e_n, e'_{n'})$ est alors donné par $p_{n,n'} = p_n p_{n'}$ car les évènements sont indépendants. Et l'entropie vérifie la propriété d'additvité :

$$S(\{p_{n,n'}\}) = S(\{p_n\}) + S(\{p'_{n'}\}) \quad (a)$$

ii) On s'intéresse à présent au cas où les évènements ne sont pas nécessairement indépendants. La probabilité d'occurence de l'évènement composé $(e_n, e'_{n'})$ est toujours notée $p_{n,n'}$, en revanche l'égalités $p_{n,n'} = p_n p_{n'}$ n'est plus vérifiée. Si on s'intéresse à un évènement du premier ensemble indépendamment de l'autre ensemble d'évènement, la probabilité d'occurence de l'évènement e_n est donné par

$$p_n = \sum_{n'=1}^{N'} p_{n,n'}$$

de même on a symétriquement :

$$p'_{n'} = \sum_{n=1}^{N} p_{n,n'}$$

On peut calculer également les probabilités conditionnelles $p_{n'}^{\prime(n)} = Proba(e'_{n'}/e_n)$

$$p_{n'}^{\prime(n)} = \frac{p_{n,n'}}{p_n}$$

L'entropie vérifie alors l'équation :

$$S(\{p_{n,n'}\}) = S(\{p_n\}) + \sum_n p_n S(\{p_{n'}^{\prime(n)}\} \quad (b)$$

Le dernier terme caractérise le manque d'information dû au fait que l'on connait les évènement $\Pi = (e_n, n = 1, \cdots, N)$ mais pas les évènements $\Pi' = (e'_{n'}, n' = 1, \cdots, N')$, on l'appelle entropie (ou information) conditionnelle de Π' connaissant Π :

$$S_\Pi(\Pi') = \sum_n p_n S(\{p_{n'}^{\prime(n)}\})$$

1.3 Quelques propriétés importantes

a) **Une inégalité utile** : en utilisant les propriétés de la fonction $\log x$, montrez que si l'on a un ensemble de d'évènements $(e_n, n = 1, \cdots, N)$ et deux distributions de probablités possibles pour ces évènements $\{p_n\}$ et $\{p'_n\}$ on a l'inégalité :

$$S(\{p_n\}) \leq -\lambda \sum_n p_n \log p'_n$$

Utilisez la formule précédente pour retrouver le postulat c).

b) **inégalité de sous additivité** : étant donné deux ensembles d'évènements $\Pi = (e_n, n = 1, \cdots, N)$ et $\Pi' = (e'_{n'}, n' = 1, \cdots, N')$, montrez que l'entropie associée aux évènements composés $\{(e_n, e'_{n'})\})$ de probabilités associées $p_{n,n'}$ vérifie l'inégalité :

$$S(\{p_{n,n'}\}) \leq S(\{p_n\}) + S(\{p_{n'}\})$$

L'égalité n'étant vérifiée que lorsque les deux ensembles d'évènements sont indépendants.
Que représente à votre avis la quantité $S(\{p_{n,n'}\}) - S(\{p_n\})$? En déduire que,

$$S_\Pi(\Pi') \leq S(\Pi')$$

c) **inégalité de concavité** : démontrez que si $\{p_n\}$ et $\{p'_n\}$ sont deux distributions de probabilités relatives au même ensemble d'évènements $\Pi = (e_n, n = 1, \cdots, N)$ et si on considère une troisième distribution obtenue par combinaison linéaire des deux premières, $p''_n = \alpha p_n + (1-\alpha) p'_n$ où $0 \leq \alpha \leq 1$ alors on a :

$$S(\{p''_n\}) \geq \alpha S(\{p_n\}) + (1-\alpha) S(\{p'_n\})$$

2 Quelques applications de l'entropie

2.1 Language et Information

On considère un alphabet à deux "lettres", point et trait comme en télégraphie, ou 0 et 1 comme en language binaire informatique. On considère G "cases", N_0 de ces cases contiennent des 0 et N_1 contiennent des 1. Toutes ces cases sont remplies (soit $G = N_0 + N_1$).
a) Ecrivez les probabilités p_0 et p_1 pour qu'une case contienne respectivement un 0 ou 1.
b) Déterminez le nombre de manières P de remplir chacune des G cases avec soit 0 soit 1, mais jamais avec les deux. En déduire l'information I (ou l'entropie, $S = I$) contenue dans un message à G symboles.
c) Si le message est long et G, N_0 et N_1 suffisamment grands, utilisez la formule de Stirling pour exprimer l'information I.
d) Exprimez $i = I/G$ en fonction de p_0 et p_1.
e) Généralisez au cas de $N_1, N_2, \cdots, N_M$ symboles distincts.

2.2 Problème de boules noires et blanches

On considère un ensemble total de N boules formé de n_+ boules noires et n_- boules blanches ($N = n_+ + n_-$). On note $M = (n_+ - n_-)$ le déséquilibre entre les boules noires et les boules blanches. Calculez l'entropie S du système en fonction de N et M.

On considère le cas où N est très grand. Développez l'expression de l'entropie à l'aide de la formule de Stirling (sauf dans les cas limites où l'un des nombres n_+ et n_- est voisin de 1, que l'on étudiera séparement) et donner une expression de S/N en fonction de $x = M/N$. Commentez la forme de la fonction $f(x) = S/N$, et discutez de sa valeur en $x = 0$.

A votre avis quel système physique pourrait se rapprocher de ce jeu de boules ?

2.3 Sac de billes et entropie de mélange

On considère deux sacs de billes contenant des billes de M couleurs différentes : le premier sac contient N billes au total, dont N_1 de la couleur (1), N_2 de la couleur $(2), \cdots, N_M$ de la couleur (M), le second sac contient N' billes au total, dont N'_1 de la couleur (1), N'_2 de la couleur $(2), \cdots, N'_M$ de la couleur (M). Comparez l'entropie du premier sac seul S_1, du deuxième sac seul S_2 et l'entropie des deux sacs mélangés S_{1+2}. (Pensez à utiliser l'inégalité de concavité avec le bon paramètre α....)

Par définition $S_{1+2} - S_1 - S_2$ est l'entropie de mélange du système. Calculez l'entropie de mélange dans le cas particulier où il n'y a que deux couleurs (1) et (2), et que le premier sac contient uniquement des billes de couleur (1) et le deuxième sac uniquement des billes de couleur (2). On exprimera $S/(N+N')$ en fonction de $c = N/(N+N')$ qui est la "concentration" de billes de couleur (1).