1 Maximisantes de l'entropie de Shannon

Rappels

Définition de l'entropie d'une loi discrète.

Pour une variable aléatoire (v.a.) X de loi de probabilité discrète $\mathcal{P} = \{p_j\}_{j\in\mathbb{N}}$, l'entropie de Shannon est définie par

$$H(X) = -\sum_{j=0}^{+\infty} p_j \log_2 p_j.$$

Par abus de notation, on écrit indifféremment $H(X)$, $H(\mathcal{P})$ ou $H(p_0, p_1, \cdots)$.

1.1 Maximum d'entropie sous contrainte de support

La v.a. X suit une loi discrète à nombre d'états N fini.

Les probabilités des états sont $\{p_0, p_1, \cdots, p_{N-1}\}$.

1. Montrer que $H(p_0, \cdots, p_{N-1}) \geq 0$. Dans quels cas a-t-on $H(p_0, \cdots, p_{N-1}) = 0$?
2. Soient $\mathcal{P} = \{p_j\}_{j=0\cdots N-1}$ et $\mathcal{Q} = \{q_j\}_{j=0\cdots N-1}$ deux lois de probabilité à N états.
   Montrer :
   $$D(\mathcal{P}||\mathcal{Q}) = \sum_{j=0}^{N-1} p_j \log_2 \frac{p_j}{q_j} \geq 0$$
   Dans quel cas a-t-on égalité ?
3. En choisissant la loi $\mathcal{Q} = \{q_j\}_{j=0\cdots N-1}$, montrer que $H(p_0, \cdots, p_{N-1}) \leqslant \log_2 N$.
4. Quelle est la loi qui maximise l'entropie de la v.a. X ?
   Quelle est l'entropie de cette loi ?

1.2 Octet d'information

On considère un couple (X_1, X_2) composé de 2 v.a. binaires à valeurs dans $\{0; 1\}$.

1. Quelle doit être la distribution des probabilités sur les configurations du couple pour qu'une v.a. à valeur couple soit la plus informative possible ?
   Quelle est alors l'entropie maximale ?
2. Pour un couple uniforme, quelle est la loi d'une composante binaire X_j ?

3. Si les composantes suivent une loi uniforme sur $\{0;1\}$, peut-on en déduire que le couple est uniforme sur ses 4 configurations ?
4. Pour un couple uniforme, montrer que deux composantes binaires X_1 et X_2 sont indépendantes.
5. Finalement, pour que le couple soit d'entropie maximale, quelles sont les conditions requises sur ses composantes ?
6. Commentaire sur l'extension à un octet (8-uplet) composé de 8 v.a. indépendantes.

Exercice complémentaire

1.3 Maximum d'entropie sous contrainte de moyenne

Le but de l'exercice est de savoir ce que devient la loi d'entropie maximale lorsque la contrainte de support est levée pour être remplacée par une contrainte sur la moyenne. On rappelle que :

— $\sum_{j=0}^{\infty} j\,\beta^j = \frac{\beta}{(1-\beta)^2}$ pour $0 < \beta < 1$.

— $\left\{q_k = \alpha\beta^k\right\}_{k\in\mathbb{N}}$ est une loi de probabilité de moyenne μ pour $\beta = \frac{\mu}{\mu+1}$ et $\alpha = \frac{1}{1+\mu}$.

1. Calculer l'entropie de la loi $\left\{q_k = \alpha\beta^k\right\}_{k\in\mathbb{N}}$ de moyenne μ.
2. Montrer que la loi de probabilité discrète $\left\{q_k = \alpha\beta^k\right\}_{k\in\mathbb{N}}$ maximise l'entropie d'une v.a. X à valeurs entières non négatives sous contrainte de moyenne :

$$\mathbb{E}X = \sum_{k=0}^{\infty} kq_k = \mu$$

3. Pour illustrer le fait que si la contrainte de support est levée pour une contrainte de moyenne, la loi uniforme n'est plus la loi qui maximise l'entropie, constuisons un exemple simple :
   (a) Parmi les lois uniformes à support inclus dans $\mathbb{N}$, quelle sont celles dont la moyenne est égale à $\mu = 1$?
   (b) Quelle est l'entropie maximale d'une loi uniforme de moyenne $\mu = 1$?
   (c) Comparer l'entropie de la loi uniforme de moyenne 1 à celle de la maximisante de même moyenne trouvée à la première question.