
L'entropie en probabilité, approche élémentaire

Notations.

- Dans tout le problème, n désigne un entier naturel non nul et I_n désigne l'ensemble des entiers compris entre 1 et n :

$$I_n = \{k \in \mathbb{N} \mid 1 \leq k \leq n\}.$$

- L'espérance d'une variable aléatoire X sera notée $\mathbb{E}(X)$.

Préliminaire.

1. Soit $h : \mathbb{R}^+ \rightarrow \mathbb{R}$ définie par $h(x) = \begin{cases} x \ln x & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$.

Démontrer que h est continue sur $\mathbb{R}^+$.

h est-elle dérivable en 0 ?

Déterminer les antécédents de 0 par h .

2. **Une propriété fondamentale.** Démontrer que pour tout $x > 0$ on a $\ln x \leq x - 1$, avec égalité si, et seulement si, $x = 1$.

Dans tout le problème, si X est une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, l'**entropie** de X est, sous réserve d'existence :

$$H(X) = - \sum_{x \in X(\Omega)} h(P(X = x))$$

En particulier, lorsque X est à valeurs dans l'ensemble fini $\{x_1, \dots, x_n\} \subset \mathbb{R}$, l'entropie de X existe toujours et vaut :

$$H(X) = - \sum_{i=1}^n h(p_i)$$

où, pour tout i dans $\{1, \dots, n\}$, $p_i = P(X = x_i)$.

Partie I. Exemples.

3. Dans cette question U_n suit la loi uniforme sur I_n . Déterminer $H(U_n)$.
4. Soit X une variable aléatoire à valeurs dans une partie finie de $\mathbb{R}$. Démontrer que $H(X) = 0$ si, et seulement si, X est presque sûrement constante.
5. Pour x dans $[0, 1]$ on pose $g(x) = -h(x) - h(1 - x)$.
Etudier g .
Soit X une variable aléatoire qui suit une loi de Bernoulli de paramètre $p \in]0, 1[$. Démontrer que $H(X) \leq \ln 2$ avec égalité si et seulement si $p = \frac{1}{2}$.
6. Soit X_1 et X_2 deux variables de Bernoulli indépendantes de paramètres respectifs p_1 et p_2 , définies sur le même espace probabilisé.
Soit Z la variable de Bernoulli telle que $P(Z = 1) = P(\text{« } X_1 + X_2 \text{ est impair »})$.
En notant $p = P(Z = 1)$, déterminer p et vérifier que $(1 - 2p) = (1 - 2p_1)(1 - 2p_2)$.
Préciser $\mathbb{E}(Z)$.

7. Soient $p \in]0, 1[$ et X une variable aléatoire suivant une loi binomiale de paramètres n et p .
 Soit Z_n la variable de Bernoulli telle que $P(Z_n = 1) = P(\text{« } X \text{ est impair »})$.
 Montrer que $1 - 2P(Z_n = 1) = (1 - 2p)^n$ (on pourra raisonner par récurrence). Montrer que $H(Z_n) \leq \ln 2$. Dans quel(s) cas a-t-on égalité ?

Partie II. Propriétés de l'entropie des variables à valeurs dans $\mathbb{N}$

Dans cette partie X_0 est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$ qui suit une loi géométrique sur $\mathbb{N}^*$ de paramètre $p \in]0, 1[$. On pose $m = \mathbb{E}(X_0)$ et, pour tout $k \in \mathbb{N}^*$, $p_k = P(X_0 = k)$. Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}^*$, $E(X) = m$ et $H(X)$ existe.
 Pour $k \in \mathbb{N}^*$, on pose $q_k = P(X = k)$ et on supposera $q_k > 0$.

8. Rappeler la valeur de m , démontrer que X_0 admet une entropie et préciser la valeur de $H(X_0)$.
 9. Démontrer que la série $\sum_{k \geq 1} q_k \ln p_k$ converge et montrer que $H(X) = - \sum_{k=1}^{+\infty} q_k \ln p_k$.
 10. Vérifier, en utilisant la question 2 du préliminaire, que pour tout $k \in \mathbb{N}^*$, on a :

$$-h(q_k) + q_k \ln p_k \leq p_k - q_k.$$

En déduire que $H(X) \leq H(X_0)$.

11. On suppose que $H(X) = H(X_0)$. Démontrer que, pour tout k dans $\mathbb{N}^*$, on a :

$$-h(q_k) + q_k \ln p_k = p_k - q_k.$$

En déduire que X suit la même loi que X_0 .

Partie III. Entropie d'un couple.

Dans cette partie, m et n sont des entiers naturels plus grand que 2, (X, Y) et (X', Y') sont deux couples de variables aléatoires discrètes définis sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$. Les variables aléatoires X et X' sont à valeurs dans I_n . Les variables aléatoires Y et Y' sont à valeurs dans I_m .

Pour tout $(i, j) \in I_n \times I_m$, on pose :

$$p_i = P(X = i), \quad q_j = P(Y = j), \quad \lambda_{i,j} = P(X = i, Y = j) \quad \text{et} \quad \mu_{i,j} = P(X' = i, Y' = j).$$

On suppose que pour tout $(i, j) \in I_n \times I_m$ on a $\lambda_{i,j} \neq 0$ et $\mu_{i,j} \neq 0$.

Deux définitions.

On définit **l'entropie du couple** (X, Y) par : $H(X, Y) = - \sum_{i=1}^n \sum_{j=1}^m h(\lambda_{i,j})$.

On définit **l'information entre les couples** (X, Y) et (X', Y') par :

$$K(X, Y, X', Y') = - \sum_{i=1}^n \sum_{j=1}^m \lambda_{i,j} \ln \frac{\mu_{i,j}}{\lambda_{i,j}}.$$

12. **Propriétés de l'information entre deux couples**

- a. Rappeler les valeurs de $\sum_{i=1}^n \sum_{j=1}^m \lambda_{i,j}$ et $\sum_{i=1}^n \sum_{j=1}^m \mu_{i,j}$.
- b. Démontrer que $K(X, Y, X', Y') = - \sum_{i=1}^n \sum_{j=1}^m \lambda_{i,j} \left(\ln \frac{\mu_{i,j}}{\lambda_{i,j}} - \frac{\mu_{i,j}}{\lambda_{i,j}} + 1 \right)$.
- c. Démontrer que $K(X, Y, X', Y') \geq 0$.
- d. Démontrer que $K(X, Y, X', Y') = 0$ si et seulement si les couples (X, Y) et (X', Y') ont même loi conjointe.
- e. On suppose ici que les variables aléatoires X' et Y' sont indépendantes et de même loi que X et Y respectivement.
- i. Démontrer que $K(X, Y, X', Y') = H(X) + H(Y) - H(X, Y)$.
- ii. En déduire que : $H(X, Y) \leq H(X) + H(Y)$ ($\clubsuit$).
Donner une CNS pour que cette inégalité soit une égalité.

On admettra, dans la suite, que l'inégalité ($\clubsuit$) reste vraie même si certains des $\lambda_{i,j}$ ou des $\mu_{i,j}$ s'annulent.

13. Entropie conditionnelle

On définit l'entropie conditionnelle de Y sachant X par : $H(Y | X) = H(X, Y) - H(X)$.
Elle mesure l'incertitude restant sur la valeur de Y sachant la valeur de X .

- a. Démontrer que $H(Y | X) \leq H(Y)$.
- b. On considère m réels $a_1, \dots, a_n$ dans $]0, 1]$.
Démontrer que pour tout i dans I_m on a :

$$\sum_{i=1}^m h(a_i) \leq h \left(\sum_{i=1}^m a_i \right) \quad (\spadesuit)$$

($\spadesuit$) est-elle encore vraie lorsque $a_1, \dots, a_n$ sont dans $[0, 1]$?

Démontrer qu'il y a égalité dans ($\spadesuit$) si et seulement s'il existe au plus un indice i dans I_m tel que $a_i \neq 0$.

- c. Démontrer que pour tout i dans I_n , $\sum_{j=1}^m h(\lambda_{i,j}) \leq h(p_i)$. En déduire que $H(Y | X) \geq 0$.
- d. Dans cette question, on suppose que $H(Y | X) = 0$. Démontrer que, pour tout $i \in I_n$ tel que $p_i > 0$, il existe un unique j dans I_m tel que $\lambda_{i,j} > 0$. On posera $j = \alpha(i)$.
Démontrer que $Y = \alpha(X)$ presque sûrement. Comment interpréter ce résultat ?

FIN DE L'ÉNONCÉ
