DS8

Un modèle probabiliste d'une expérience aléatoire représente dans un certain sens le désordre qui intervient dans l'expérience et il est donc naturel que des outils soient introduits qui permettent de mesurer l'intensité de ce désordre. C'est le cas de la notion d'entropie qui fait l'objet du présent problème. On considèrera différentes situations et notamment la façon dont on mesure l'information que deux variables aléatoires s'apportent mutuellement.
Dans la première partie on étudie le cas plus simple techniquement de variables dont la loi admet une densité. Les deuxièmes et troisièmes parties sont consacrées au cas discret. Dans la deuxième partie, on introduit les différentes notions d'entropie pour le cas de variables discrètes et dans la troisième partie, on examine comment on peut mesurer l'information apportée mutuellement par deux variables aléatoires.
Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$. Pour toute variable aléatoire Y, on notera $\mathbb{E}(Y)$ son espérance lorsqu'elle existe.

Première partie : Entropie différentielle d'une variable à densité

1. La fonction logarithme de base 2, notée $\log_2$, est définie sur $\mathbb{R}_+^*$ par : $\log_2(x) = \dfrac{\ln(x)}{\ln(2)}$.

a) Montrer que pour tout (x, y) élément de $\mathbb{R}_+^* \times \mathbb{R}_+^*$, on a : $\log_2(x\,y) = \log_2(x) + \log_2(y)$.

b) Vérifier que pour tout réel α : $\log_2(2^\alpha) = \alpha$.

c) Montrer que la fonction $\log_2$ est concave sur $\mathbb{R}_+^*$.

2. Soit X une variable aléatoire réelle à densité, et soit f une densité de X. On appelle **support** de f l'ensemble $I = \{x \in \mathbb{R} \mid f(x) > 0\}$, et on suppose que I est un intervalle de $\mathbb{R}$ d'extrémités a et b ($a < b$, a et b finis ou infinis). L'**entropie différentielle** de X est, sous réserve d'existence, le réel :

$$h(X) \;=\; -\int_a^b f(x)\,\log_2\big(f(x)\big)\,dx$$

Montrer : $h(X) = -\mathbb{E}\Big(\log_2\big(f(X)\big)\Big)$.

3. Soit X une variable aléatoire de densité f de support I, intervalle de $\mathbb{R}$ d'extrémités a et b.
On suppose que X admet une entropie différentielle.

a) Soit c un réel, et soit Y la variable aléatoire définie par $Y = c + X$.

(i) Déterminer une densité de Y.

(ii) Justifier l'existence de l'entropie différentielle $h(Y)$, et la déterminer en fonction de $h(X)$.

b) Soit α un réel strictement positif, et soit Z la variable aléatoire définie par $Z = \alpha\,X$.

(i) Déterminer une densité de Z.

(ii) Justifier l'existence de l'entropie différentielle $h(Z)$, et la déterminer en fonction de $h(X)$.

4. On détermine dans cette question l'entropie différentielle de quelques variables aléatoires suivant des lois classiques.

a) Soit $a > 0$. On considère X une variable aléatoire de loi uniforme sur $[0, a]$.

(i) Donner une densité de X.

(ii) Justifier l'existence de l'entropie différentielle $h(X)$, et la déterminer.

(iii) Déterminer une condition nécessaire et suffisante sur a pour que $h(X) > 0$.

b) On considère Y une variable aléatoire de loi normale centrée réduite. Montrer que Y admet une entropie différentielle et : $h(Y) = \dfrac{1}{2}\ \log_2(2\pi \mathrm{e})$.

c) On considère Z une variable aléatoire de loi exponentielle de paramètre λ ($\lambda > 0$).
Justifier l'existence de l'entropie différentielle $h(Z)$ et la déterminer.

d) Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = \dfrac{1}{2}\ \lambda\, \mathrm{e}^{-\lambda\,|x|}$ ($\lambda > 0$).

(i) Montrer que f est une densité de probabilité sur $\mathbb{R}$.

(ii) Soit W une variable aléatoire de densité f.
Justifier l'existence de l'entropie différentielle $h(W)$ et la déterminer.

5. On dit qu'un couple (X,Y) de variables aléatoires est un couple gaussien centré si, pour tout $(\alpha,\beta) \in \mathbb{R}^2$, $\alpha\, X + \beta\, Y$ est une variable de loi normale centrée, c'est-à-dire qu'il existe $\gamma \in \mathbb{R}$ et une variable aléatoire Z de loi normale centrée réduite tels que $\alpha\, X + \beta\, Y$ a même loi que $\gamma\, Z$. On considère un tel couple (X,Y) et on noté σ^2 la variance de X. On suppose : $\sigma^2 > 0$.

a) Montrer que X suit une loi normale centrée.

b) Calculer $h(X)$.

c) On suppose désormais que X et Y suivent la même loi normale centrée de variance σ^2 et on admet que les propriétés de l'espérance des variables discrètes se généralisent aux variables aléatoires quelconques.

(i) Montrer que $\mathbb{E}(XY)$ existe.

(ii) Montrer de plus, pour tout réel λ : $\lambda^2\, \mathbb{E}\left(Y^2\right) + 2\lambda\, \mathbb{E}(XY) + \mathbb{E}\left(X^2\right)\ \geqslant\ 0$.

(iii) En déduire : $\big(\mathbb{E}(XY)\big)^2\ \leqslant\ \mathbb{E}\left(X^2\right)\, \mathbb{E}\left(Y^2\right)$.

(iv) On pose $\rho = \dfrac{\mathbb{E}(XY)}{\sigma^2}$. Montrer : $\rho \in [-1,1]$.

(v) Que vaut ρ si X et Y sont indépendantes ?

d) On suppose $|\rho| < 1$. On appelle **entropie jointe** du couple (X,Y) le réel :

$$h(X,Y)\ =\ \log_2\left(2\pi\, \mathrm{e}\, \sigma^2\, \sqrt{1-\rho^2}\right)$$

(i) À quelle condition $h(X,Y)$ est-elle nulle ?

(ii) L'**information mutuelle** de X et Y est définie par :

$$I(X,Y)\ =\ h(X) + h(Y) - h(X,Y)$$

Calculer $I(X,Y)$.

(iii) Montrer : $I(X,Y) \geqslant 0$.

(iv) Quelle est la limite de $I(X,Y)$ quand ρ tend vers 1 ?

Deuxième partie : Généralités sur l'entropie des variables discrètes

Soit A un ensemble fini non vide. On dit que X est une variable aléatoire dont la loi est à support A, si X est à valeurs dans A et si pour tout $x \in A$: $\mathbb{P}([X = x]) > 0$.

6. Soit X une variable aléatoire de loi à support $\{0, 1, 2, \ldots, n\}$ où n est un entier naturel. On appelle **entropie** de X le réel :

$$H(X) \ = \ -\sum_{k=0}^{n} \mathbb{P}([X = k]) \log_2 \big(\mathbb{P}([X = k])\big)$$

a) On définit la fonction $g : \{0, \ldots, n\} \to \mathbb{R}$ en posant $g(k) = \log_2 \big(\mathbb{P}([X = k])\big)$ pour k élément de $\{0, 1, \ldots, n\}$. Montrer : $H(X) = -\mathbb{E}\big(g(X)\big)$.

b) Montrer : $H(X) \geqslant 0$.

c) Soit p un réel tel que $0 < p < 1$.
On suppose dans cette question que X suit la loi de Bernoulli $\mathcal{B}(p)$.

(i) Calculer $H(X)$ en fonction de p. On note ψ la fonction qui, à p, associe $H(X)$.

(ii) Montrer que ψ est concave sur $]0, 1[$.

(iii) Déterminer la valeur p_0 où ψ est maximale.

d) On suppose dans cette question que la loi de X est à support $\{0, 1, 2, 3\}$ avec les probabilités :

$$\mathbb{P}([X = 0]) = \frac{1}{2} \quad ; \quad \mathbb{P}([X = 1]) = \frac{1}{4} \quad ; \quad \mathbb{P}([X = 2]) = \mathbb{P}([X = 3]) = \frac{1}{8}$$

Calculer $H(X)$.

7. On souhaite écrire une fonction en **Python** pour calculer l'entropie d'une variable aléatoire X dont le support de la loi est de la forme $A = \{0, 1, \ldots, n\}$ où n est un entier naturel. On suppose que la liste `P` de **Python** est telke que pour tout k de A, `P[k]` $= \mathbb{P}([X = k])$. Compléter la fonction ci-dessous d'argument `P` qui renvoie l'entropie de X, c'est-à-dire $-\sum_{k=0}^{n} \mathbb{P}([X = k]) \log_2 \big(\mathbb{P}([X = k])\big)$.

```python
def Entropie(P) :
    ...
    return h
```

Si nécessaire, on pourra utiliser l'instruction `len(P)` qui donne le nombre d'éléments de `P`.
On souhaite maintenant démontrer quelques inégalités concernant l'entropie.

8. On commence par une inégalité générale, appelée **Inégalité de Jensen**.

a) Soit $N \geqslant 2$. Soit X une variable aléatoire de loi à support $\{x_1, x_2, \ldots, x_N\}$ où les x_i sont des éléments distincts de $\mathbb{R}_+$. On pose $\mathbb{P}([X = x_i]) = p_i$.
Montrer que, pour tout $1 \leqslant i \leqslant N$, on a : $p_i < 1$.

On désire démontrer par récurrence la propriété suivante :

$\mathcal{P}(N)$: **Pour toute fonction φ convexe sur $\mathbb{R}_+$, si X est une variable aléatoire de loi à support $A \subset \mathbb{R}_+$ avec $\mathrm{Card}(A) = N$, on a :** $\mathbb{E}\big(\varphi(X)\big) \geqslant \varphi\big(\mathbb{E}(X)\big)$

b) Montrer que $\mathcal{P}(2)$ est vraie.

c) Soit $N \geqslant 3$. On suppose que $\mathcal{P}(N-1)$ est vérifiée. Soit X une variable aléatoire de loi à support $A = \{x_1, x_2, \dots, x_N\}$ où les x_i sont des éléments distincts de $\mathbb{R}_+$. On pose : $\mathbb{P}([X = x_i]) = p_i$. Pour i tel que $1 \leqslant i \leqslant N-1$, on pose : $p'_i = \dfrac{p_i}{1-p_N}$.

(i) Montrer : $\sum\limits_{i=1}^{N-1} p'_i = 1$ et $\forall i \in [\![1, N-1]\!]$, $0 < p'_i < 1$.

(ii) Soit Y une variable aléatoire de loi à support $\{x_1, \dots, x_{N-1}\}$ telle que $\mathbb{P}([Y = x_i]) = p'_i$ pour $1 \leqslant i \leqslant N-1$. Montrer : $\sum\limits_{i=1}^{N-1} p'_i\, \varphi(x_i) \;\geqslant\; \varphi\left(\sum\limits_{i=1}^{N-1} p'_i\, x_i\right)$.

(iii) Montrer : $\mathbb{E}\big(\varphi(X)\big) \;\geqslant\; \varphi\big(\mathbb{E}(X)\big)$.

d) Montrer que, si φ est *concave* sur $\mathbb{R}_+$, on a : $\mathbb{E}\big(\varphi(X)\big) \;\leqslant\; \varphi\big(\mathbb{E}(X)\big)$.

9. Soit X une variable aléatoire de loi à support $\{0, 1, \dots, n\}$.
On pose, pour k tel que $0 \leqslant k \leqslant n$, $p_k = \mathbb{P}([X = k])$.

a) Montrer : $\sum\limits_{k=0}^{n} p_k \log_2\left(\dfrac{1}{(n+1)\, p_k}\right) \;\leqslant\; \log_2\left(\sum\limits_{k=0}^{n} \dfrac{p_k}{(n+1)\, p_k}\right) \;=\; 0.$

b) Montrer : $\sum\limits_{k=0}^{n} p_k \log_2\big((n+1)\, p_k\big) \;=\; \log_2(n+1) - H(X)$.

c) Montrer : $H(X) \leqslant \log_2(n+1)$.

d) On suppose que X suit la loi uniforme sur $\{0, 1, \dots, N\}$. Calculer $H(X)$.

10. Soient X et Y deux variables aléatoires *de même loi* à support $\{0, 1, \dots, n\}$.
On suppose en outre X et Y indépendantes.

a) Montrer : $\mathbb{P}([X = Y]) \;=\; \sum\limits_{k=0}^{n} \big(\mathbb{P}([X = k])\big)^2$.

b) On pose $v(k) = \mathbb{P}([X = k])$ pour tout k élément de $\{0, 1, \dots, n\}$. Montrer :

$$2^{\mathbb{E}\left(\log_2\left(v(X)\right)\right)} \;\leqslant\; \mathbb{E}\left(2^{\log_2\left(v(X)\right)}\right) \;=\; \mathbb{E}\big(v(X)\big)$$

c) En déduire : $2^{-H(X)} \;\leqslant\; \mathbb{P}([X = Y])$.

d) Donner un exemple de loi où l'inégalité précédente est une égalité.

Troisième partie : Entropie jointe et information mutuelle de deux variables discrètes

Soient X et Y deux variables aléatoires de lois à support $\{0, 1, \dots, n\}$.
On appelle **entropie jointe** de X et Y le réel :

$$H(X,Y) \;=\; -\sum_{k=0}^{n} \sum_{j=0}^{n} \mathbb{P}([X = k] \cap [Y = j]) \log_2\big(\mathbb{P}([X = k] \cap [Y = j])\big)$$

avec la convention : $0 \times \log_2(0) = 0$.

11. ***a)*** On définit la fonction $g : \{0, 1, \dots, n\}^2 \to \mathbb{R} \cup \{-\infty\}$ en posant pour $(k, j) \in \{0, 1, \dots, n\}^2$:

$$g(k,j) \;=\; \log_2\big(\mathbb{P}([X = k] \cap [Y = j])\big)$$

Montrer : $H(X,Y) = -\mathbb{E}\big(g(X,Y)\big)$.

b) Montrer : $H(X,Y) = H(Y,X)$.

c) Pour tout k tel que $0 \leqslant k \leqslant n$, on pose :

$$H(Y \mid X = k) \;=\; -\sum_{j=0}^{n} \mathbb{P}_{[X=k]}([Y = j]) \log_2 \left(\mathbb{P}_{[X=k]}([Y = j])\right)$$

On appelle **entropie conditionnelle** de Y sachant X le réel :

$$H(Y \mid X) \;=\; \sum_{k=0}^{n} \mathbb{P}([X = k])\, H(Y \mid X = k)$$

Montrer : $H(X,Y) \;=\; H(X) + H(Y \mid X)$.

d) Montrer que pour tout couple de variables aléatoires X et Y de lois à support $\{0, 1, \ldots, n\}$, on a :

$$H(X) - H(X \mid Y) \;=\; H(Y) - H(Y \mid X)$$

12. On considère dans cette question deux variables aléatoires de lois à support $\{0, 1, 2, 3\}$. On suppose que la loi conjointe de (X, Y) est donnée par le tableau suivant :

j \ k	0	1	2	3
0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
1	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
3	$\frac{1}{4}$	0	0	0

(on lit dans la $k^{\text{ème}}$ colonne et la $j^{\text{ème}}$ ligne la valeur de $\mathbb{P}([X = k] \cap [Y = j])$)

a) Déterminer la loi de X et montrer : $H(X) = \dfrac{7}{4}$.

b) Déterminer la loi de Y et calculer $H(Y)$.

c) Montrer : $H(X \mid Y) = \dfrac{11}{8}$.

d) Que vaut $H(Y \mid X)$?

e) Calculer $H(X,Y)$.

13. Soient X et Y deux variables aléatoires de lois à support $\{0, 1, \ldots, n\}$. On appelle **information mutuelle** de X et de Y le réel :

$$I(X,Y) \;=\; \sum_{k=0}^{n} \sum_{j=0}^{n} \mathbb{P}([X = k] \cap [Y = j]) \log_2 \left(\frac{\mathbb{P}([X = k] \cap [Y = j])}{\mathbb{P}([X = k])\, \mathbb{P}([Y = j])}\right)$$

a) Montrer : $I(X,Y) = I(Y,X)$.

b) Montrer : $I(X,Y) = H(X) - H(X \mid Y)$.

c) Montrer : $I(X,X) = H(X)$.

d) Que vaut $I(X,Y)$ si X et Y sont indépendantes ?

14. Soient X et Y deux variables aléatoires de lois à support $\{0, 1, \ldots, n\}$. On fixe $0 \leqslant k \leqslant n$.

Pour $0 \leqslant j \leqslant n$, on pose : $p_j = \dfrac{\mathbb{P}([X = k] \cap [Y = j])}{\mathbb{P}([X = k])}$.

On suppose que $p_j > 0$ pour tout $0 \leqslant j \leqslant n$ et on pose : $x_j = \dfrac{\mathbb{P}([X = k])\, \mathbb{P}([Y = j])}{\mathbb{P}([X = k] \cap [Y = j])}$.

a) Montrer : $\displaystyle\sum_{j=0}^{n} p_j = 1$.

b) Soit Z_k une variable aléatoire de loi à support $\{x_0, \ldots, x_n\}$ dont la loi est donnée par $\mathbb{P}([Z_k = x_j]) = p_j$ pour $0 \leqslant j \leqslant n$. Montrer :

$$\mathbb{E}\big(\log_2(Z_k)\big) \;\leqslant\; 0$$

c) En déduire : $I(X, Y) \geqslant 0$.