Devoir à la maison n°14

La notation tiendra particulièrement compte de **la qualité de la rédaction, la <u>précision</u> des raisonnements et l'énoncé des <u>formules utilisées</u>**.

Problème : Entropie et probabilité

$(\Omega, \mathcal{A}, \mathbf{P})$ désigne un espace probabilisé.

— On note φ la fonction définie sur $]0,1]$ par $\varphi : x \mapsto -\ln(x)$
Pour un événement A de probabilité non nulle, on pose $i(A) = \varphi(\mathbf{P}(A))$.

— On note h la fonction définie sur $\mathbb{R}_+$ par

$$h(0) = 0 \text{ si} x \geqslant 0 \qquad \text{et pour } x > 0, \quad h(x) = -x\ln(x)$$

— Pour une variable aléatoire X discrète définie sur $(\Omega, \mathcal{A}, P)$ à valeurs réelles, on pose sous réserve d'existence :

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

Si X est à valeurs dans un ensemble fini, alors $H(X)$ existe et la somme précédente est finie.

A. Exemples. (Incertitude d'événements et entropie de variable aléatoire)

1. On choisit une carte au hasard dans un jeu de 32 cartes.
Soit A l'événement ≪ la carte tirée est la dame de cœur ≫.
Que valent $\mathbf{P}(A)$ et $i(A)$?
2. Soit $n \in \mathbb{N}^*$. On lance n fois une pièce équilibrée. B est l'événement ≪ obtenir n fois PILE ≫.
Préciser $i(B)$.
3. Vérifier les points suivants (*) :
— Pour un événement Ω' (quasi-)certain : $i(\Omega') = 0$.
— Si A et l'événement contraire $\overline{A}$ sont équiprobables, alors $i(A) = \ln 2$.
— Si A et B sont indépendants pour la probabilité $\mathbf{P}$ et si $\mathbf{P}(A \cap B) \neq 0$, alors $i(A \cap B) = i(A) + i(B)$.
4. Préciser $i(A_1 \cap A_2 \cap \cdots \cap A_n)$ quand les événements $A_1, A_2, \ldots, A_n$ sont mutuellement indépendants et $\mathbf{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \neq 0$.
En déduire une nouvelle démonstration de A.2.
5. Soit A et B deux événements tels que $A \subset B$ et $P(A) \neq 0$.
Comparer $i(A)$ et $i(B)$.
6. Soit U_3, une variable aléatoire qui suit la loi uniforme sur $\{1, 2, 3\}$, que vaut $\mathbf{H}(U_3)$?
On voit ici que H est une sorte d'incertitude moyenne pour la variable X.
7. Si on suppose $\mathbf{P}(Z = 1) = 1/4, \mathbf{P}(Z = 2) = 1/4$ et $\mathbf{P}(Z = 3) = 1/2$, que vaut $\mathbf{H}(Z)$?
Comparer $\mathbf{H}(Z)$ et $\mathbf{H}(U_3)$.

B. Étude de la fonction h

1. Démontrer que h est continue sur $\mathbb{R}_+$. h est-elle de classe $\mathcal{C}^1$ sur $\mathbb{R} - +$?
2. Justifier l'inégalité suivante :
$$(\mathcal{I}) \qquad \forall\, u \in \mathbb{R}_+, h(u) \leqslant 1 - u$$
avec égalité si et seulement si $u = 1$
3. Faire la représentation graphique de h et $u \mapsto u - 1$. On prendra $4cm$ pour unité.
4. On cherche à montrer que φ est ≪ la seule ≫fonction qui vérifie les hypothèses (*).
On considère ψ continue sur $]0,1]$ et telle que $\psi(1) = 0$, $\psi(\frac{1}{2}) = a$ et $\forall\, p, q \in]0,1]$, $\psi(pq) = \psi(p) + \psi(q)$.
(a) Montrer à l'aide d'un changement de variable affine :
$$\forall\, p \in]0,1], \qquad \frac{1}{p}\int_{p/2}^{p} \psi(t)\mathrm{d}t = \frac{1}{2}\psi(p) + \int_{1/2}^{1} \psi(q)\mathrm{d}q$$
(b) En déduire que ψ est dérivable sur $]0,1]$ et en dérivant $p \mapsto p\psi(p)$, démontrer que :
$$\forall\, p \in]0,1], \qquad -\frac{a}{2} = \frac{1}{2}p\psi'(p) + \int_{1/2}^{1} \psi(q)\mathrm{d}q$$
(c) En déduire qu'il existe α et β telle que $\psi : t \mapsto \alpha \ln x + \beta$

C. Entropie des variables aléatoires discrètes d'univers image fini

Soient $n \in \mathbb{N}$ et X une variable aléatoire définie sur $(\Omega, \mathcal{T}, \mathbf{P})$ à valeurs dans $X(\Omega) = \{x_1, x_2, \dots x_n\}$.

1. Soit B_p, la variable de Bernoulli de paramétre $p \in]0,1[$, de variance $\mathbf{V}(B_p)$ et d'entropie $\mathbf{H}(B_p)$. Soit ψ la fonction définie par :

$$\forall\ p \in]0,1[, \qquad \psi(p) = \mathbf{H}(B_p) - \mathbf{V}(B_p)$$

   (a) Montrer que l'expression de $\psi(p)$ en fonction de p est donnée par

$$\forall\ p \in]0,1[\quad \psi(p) = -p\ln(p) + (p-1)[p + \ln(1-p)]$$

   (b) Calculer la dérivée ψ' de ψ

   (c) Calculer la dérivée seconde ψ'' de ψ et étudier les variations de ψ'.
   Calculer $\psi'(\frac{1}{2})$

   (d) En déduire les variations de ψ

   (e) Déterminer les limites suivantes : $\lim\limits_{p \to 0^+} \psi(p)$ et $\lim\limits_{p \to 1^-} \psi(p)$.
   Prouver que pour tout réel $p \in]0,1[$, on a $\mathbf{V}(B_p) \leqslant \mathbf{H}(B_p)$.

2. Soit n un entier non nul et U_n la variable uniforme sur $\mathbb{N}_n$.

   (a) Vérifier que : $\mathbf{H}(U_n) = \ln n$

   (b) Rappeler l'expression de la variance $\mathbf{V}(U_n)$ de U_n et établir que

$$\mathbf{H}(U_n) = \frac{1}{2}\ln(12\mathbf{V}(U_n) + 1)$$

   (c) En déduire que l'entropie et la variance de variables uniformes sont simultanément croissantes.

3. Montrer , en utilisant l'inégalité $(\mathcal{I})$ pour les $u = np_i$, que pour toute variable aléatoire discréte X, à valeurs dans $X(\Omega) = \{x_1, x_2 \dots x_n\}$, telle que $p_i = \mathbf{P}(X = x_i)$

$$\mathbf{H}(X) \leqslant \mathbf{H}(U_n)$$

   avec égalité si et seulement si $X = U_n$.
   On a montré que parmi les variables aléatoires discrétes X, la loi uniforme U_n réalise le maximum de l'entropie.

D. Entropie et distribution de Boltzman

On considère une variable aléatoire X à valeur dans $\{E_1, E_2, \dots E_n\}$.
On cherche la distribution p_x des probabilités à donner à X de sorte que :
— $\sum p_x = 1$
— $\mathbf{E}(X) = E$, c'est une donnée connue du problème
— $H(X)$ soit maximale

Il s'agit d'un problème d'optimisation (de H) sous contraintes (2 contraintes : la somme des probabilités et l'espérance).
Dans ce cas la stratégie mathématiques est la suivante : on ajoute deux variables supplémentaires aux problèmes, on crée ainsi une nouvelle fonction que l'on optimise par un calcul de point critique (les dérivées sont nulles). C'est la méthode des multiplicateurs de Lagrange

Soit $F : (p_1, p_2, \dots, p_n, \lambda_1, \lambda_2) \longmapsto -\sum\limits_{i=1}^{n} h(p_i) + \lambda_1(\sum\limits_{i=1}^{n} p_i - 1) + \lambda_2(\sum\limits_{i=1}^{n} E_i p_i - E)$.

1. Calculer $\dfrac{\partial F}{\partial p_i}(p_1, \dots, p_n, \lambda_1, \lambda_2)$, $\dfrac{\partial F}{\partial \lambda_i}(p_1, \dots, p_n, \lambda_1, \lambda_2)$.

2. Résoudre le système

$$\begin{cases} \forall\ i \in [\![1,n]\!], \dfrac{\partial F}{\partial p_i}(p_1, \dots, p_n, \lambda_1, \lambda_2) = 0 \\ \forall\ i \in [\![1,2]\!], \dfrac{\partial F}{\partial \lambda_i}(p_1, \dots, p_n, \lambda_1, \lambda_2) = 0 \end{cases}$$

   Le système n'est pas linéaire, on pourra exploiter la substitution.

3. En déduire la distribution optimale de probabilité à donner à chaque p_k (sous forme d'équation)

4. Quel rapport avec la distribution de Boltzmann ?

L'objectif du problème est d'étudier les rudiments de la théorie de la communication *- ou théorie de l'information - introduite en 1948 par Claude Shannon.*
En théorie de l'information, $i(A)$ est appelé incertitude de l'événement *A et $H(X)$ est l'*incertitude moyenne *- ou* entropie *- de X.*