Devoir surveillé n°10

Durée de l'épreuve : 4 heures
La calculatrice est interdite

Le devoir est composé d'un problème.
Lorsqu'une question est jugée, a priori, plus difficile, elle est précédée du symbole $(*)$ voire $(**)$.
La notation tiendra particulièrement compte de **la qualité de la rédaction, la précision des raisonnements et l'énoncé des formules utilisées**.

BON COURAGE

Problème - Surprise !

Objectifs

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

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

— On note φ la fonction définie sur $\mathbb{R}$ par

$$\varphi(x) = 0 \text{ si } x \leqslant 0 \qquad \text{et pour } x > 0, \quad \varphi(x) = -x\ln(x) = x \times L(x)$$

— Pour une variable aléatoire X discrète définie sur $(\Omega, \mathcal{A}, P)$ à valeurs réelles, on définit son entropie par (sous réserve d'existence - *cas infini*) :

$$\mathbf{S}(X) = \sum_{x \in X(\Omega)} \varphi(\mathbf{P}(X = x))$$

Si X est à valeurs dans un ensemble fini, alors $\mathbf{S}(X)$ existe et la somme précédente est finie.
On note que $\mathbf{S}(X)$ *ne dépend que de* $p_x (= \mathbf{P}(X = x))$ *et non de* $X(\Omega)$.

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

On étudie quelques cas particuliers.

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 $s(A)$?
2. Soit $n \in \mathbb{N}^*$. On lance n fois une pièce équilibrée. B est l'événement « obtenir n fois PILE ».
Préciser $s(B)$.
3. Vérifier les points suivants :
— Pour un événement Ω' (quasi-)certain : $s(\Omega') = 0$.
— Si A et l'événement contraire $\overline{A}$ sont équiprobables, alors $s(A) = \ln 2$.
— Si A et B sont indépendants pour la probabilité $\mathbf{P}$ et si $\mathbf{P}(A \cap B) \neq 0$, alors $s(A \cap B) = s(A) + s(B)$.
4. Préciser $s(A_1 \cap A_2 \cap \cdots \cap A_n)$ quand les événements $A_1, A_2, \ldots, A_n$ sont mutuellement indépendants et $P(A_1 \cap A_2 \cap \cdots \cap A_n) \neq 0$.
En déduire une nouvelle démonstration de 2.
5. Soit A et B deux événements tels que $A \subset B$ et $\mathbf{P}(A) \neq 0$.
Comparer $s(A)$ et $s(B)$.
6. Soit U_3, une variable aléatoire qui suit la loi uniforme sur $\{1, 2, 3\}$, que vaut $\mathbf{S}(U_3)$?
7. Soit Z une nouvelle variable aléatoire telle que $Z(\Omega) = \{1, 2, 3\}$ également.
On suppose $\mathbf{P}(Z = 1) = 1/4, \mathbf{P}(Z = 2) = 1/4$ et $\mathbf{P}(Z = 3) = 1/2$, que vaut $\mathbf{S}(Z)$?
Comparer $\mathbf{S}(Z)$ et $\mathbf{S}(U_3)$.

Partie II. Mathématisation de la surprise

On montre que

1. Démontrer que φ est continue sur $\mathbb{R}$. φ est-elle de classe $\mathcal{C}^1$ sur $[0,1]$?
2. Justifier l'inégalité suivante :

$$\forall\, u \in \mathbb{R}, \varphi(u) \leqslant 1-u \qquad (\mathcal{I}_1)$$

avec égalité si et seulement si $u = 1$

3. Faire la représentation graphique de φ et $u \mapsto u-1$.
4. On cherche à montrer que s est « la seule »fonction qui vérifie les hypothèses (*) suivantes :

$$(*)\begin{cases} (i) & f(1)=0 \\ (ii) & f \text{ décroissante sur }]0,1] \\ (iii) & \forall\, p,q \in]0,1]^2,\ f(pq) = f(p)+f(q) \\ (iv) & f \text{ continue sur }]0,1] \end{cases}$$

On considère donc ψ qui vérifie $(*)$.

(a) Montrer que chacune des conditions de $(*)$ modélise l'effet de surprise d'un événement A lorsqu'elle s'applique à la probabilité $\mathbf{P}(A)$

(b) Montrer à l'aide d'un changement de variable affine :

$$\forall\, p \in]0,1], \qquad \frac{1}{p}\int_{p/2}^{p} \psi(t)\mathrm{d}t = \frac{1}{2}\psi(p) + \int_{1/2}^{1} \psi(q)\mathrm{d}q$$

(c) En déduire que ψ est dérivable sur $]0,1]$ et en dérivant $p \mapsto p\psi(p)$, démontrer que :

$$\exists\, a \in \mathbb{R} \text{ tel que } \forall\, p \in]0,1], \qquad -\frac{a}{2} = \frac{1}{2}p\psi'(p) + \int_{1/2}^{1} \psi(q)\mathrm{d}q$$

(d) En déduire qu'il existe $\alpha > 0$ telle que $\psi : x \mapsto \alpha L(x)$

(e) Que vaut $\lim\limits_{x \to 0^+} \psi(x)$? Interprétez ce résultat.

III. Entropie des variables aléatoires discrètes d'univers image fini puis dénombrable

Soient $n \in \mathbb{N}$ et X une variable aléatoire définie sur $(\Omega, \mathcal{T}, \mathbf{P})$ à valeurs dans $X(\Omega) = \{x_1, x_2, \ldots x_n\}$. On rappelle que l'entropie de X est le réel $\mathbf{S}(X)$ défini par

$$\mathbf{S}(X) = -\sum_{i=1}^{n} p_i \ln(p_i) \qquad \text{où } p_i = \mathbf{P}(X = x_i) > 0$$

1. Exprimer $\mathbf{S}(X)$ comme une espérance, puis en terme de surprise moyenne de la variable X.
2. Montrer que $\mathbf{S}(X) \geqslant 0$ et que $\mathbf{S}(X) = 0$ si et seulement si X est une variable aléatoire certaine.
3. Soit n un entier non nul et U_n la variable uniforme sur $\mathbb{N}_n$.

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

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

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

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

4. Montrer , en utilisant l'inégalité $(\mathcal{I}_1)$ pour les $u = np_i$, que pour toute variable aléatoire X, à valeurs dans $X(\Omega) = \{x_1, x_2 \ldots x_n\}$, telle $p_i = \mathbf{P}(X = x_i)$

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

avec égalité si et seulement si $X = U_n$.

On a montré que parmi les variables aléatoires finie X, la loi uniforme U_n réalise le maximum de l'entropie.

Partie IV. Entropie, distribution de Boltzman et loi normale

L'entropie mesure la moyenne de la surprise d'une variable aléatoire ou bien on peut la considérer comme la mesure du *manque d'information* par rapport à la certitude complète, d'entropie nulle (événement certain).

La loi la plus légitime est donc celle qui ne suppose pas trop d'information, celle dont l'entropie est maximale. La fin de la partie précédente montre que lorsqu'on ne connait que $X(\Omega)$ fini, la distribution de probabilité *la plus légitime* est la distribution uniforme.

Supposons que l'on a un dé à six faces, dont on sait que la moyenne des résultats des lancers est $E \neq \frac{7}{2}$.

On cherche à savoir quelle est la loi de probabilité à associer à ce dé, la plus légitime.

On considère une variable aléatoire X à valeur dans $[\![1, 6]\!]$.

On cherche la distribution $p_i = \mathbf{P}(X = i)$ des probabilités à donner à X de sorte que :

$$\left\{\begin{array}{l} - \quad \displaystyle\sum_{i=1}^{6} p_i = 1 \\ - \quad \mathbf{E}(X) = \displaystyle\sum_{i=1}^{6} ip_i = E, \text{ où } E \text{ est une donnée connue du problème} \\ - \quad \mathbf{S}(X) = -\displaystyle\sum_{i=1}^{6} p_i \ln p_i \text{ soit maximale} \end{array}\right.$$

1. On note $\mathcal{A} = \{\overrightarrow{p} = (p_1, \dots p_6) \mid \sum_{i=1}^{6} p_i = 1 \text{ et} \sum_{i=1}^{6} ip_i = E\}$.

   Montrer que si $\overrightarrow{p} = (p_1, \dots p_6) \in \mathcal{A}$, alors il existe $(a_{i,j})_{i\in\{1,2\}, j\in\{3,\dots 6\}}$, α_1 et α_2 tel que

   $$\forall\ i \in \{1, 2\}, \qquad p_i = \sum_{j=3}^{6} a_{i,j} p_j + \alpha_i$$

   En déduire la structure algébrique de l'ensemble $\mathcal{A}$.

   On donnera les valeurs explicites des nombres $a_{i,j}$ et α_i

2. On considère alors

$$\begin{array}{rccl} f: & \mathbb{R}^4 & \to & \mathbb{R} \\ & (p_3, \dots p_6) & \mapsto & \mathbf{S}(X) = -\left(\sum_{j=3}^{6} a_{1,j}p_j + \alpha_1\right) \ln \left(\sum_{j=3}^{6} a_{1,j}p_j + \alpha_1\right) - \left(\sum_{j=3}^{6} a_{2,j}p_j + \alpha_2\right) \ln \left(\sum_{j=3}^{6} a_{2,j}p_j + \alpha_2\right) \\ & & & - p_3 \ln p_3 - p_4 \ln p_4 - p_5 \ln p_5 - p_6 \ln p_6 \end{array}$$

   > *On admet que si $\mathbf{S}(X)$ est maximale sur $\mathcal{A}$ en $\overrightarrow{p}^* = (p_1^*, \dots p_6^*)$,*
   > *alors pour tout $j \in [\![3, 6]\!]$, $\dfrac{\partial f}{\partial p_j}(\overrightarrow{p}^*) = 0$ (calcul de dérivation partielle, vu en physique).*

   Montrer que pour $\forall\ \overrightarrow{p} \in \mathcal{A}$, $\forall\ j \geqslant [\![3, 6]\!]$, $-\dfrac{\partial f}{\partial p_j}(\overrightarrow{p}) = \ln p_j + 1 + a_{1,j}(\ln p_1 + 1) + a_{2,j}(\ln p_2 + 1)$.

3. En déduire qu'il existe A et B (indépendante de i) tels que pour tout $j \in [\![1, 6]\!]$, $p_j^* = \exp(Aj+B)$.
4. Que pensez-vous de la distribution de probabilités ainsi obtenue ?
5. (*) *Question à peu de points, facile mais longue, à faire que s'il vous reste du temps...*

   On considère un nouveau problème.

   Supposons que l'on sait que $X(\Omega) = [\![-n, n]\!]$, $\mathbf{E}(X) = 0$ et $\mathbf{V}(X) = 1$.

   Quelle est la distribution la plus légitime pour X ?

Partie V. Entropie de couples de variable aléatoire

On définit l'entropie de couple de variables aléatoires finies (X, Y) par

$$\mathbf{S}(X, Y) = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} \varphi(\mathbf{P}(X = x, Y = y))$$

Si $X(\Omega) = X'(\Omega)$ et $Y(\Omega) = Y'(\Omega)$, on définie l'information entre les couples (X, Y) et (X', Y') par

$$K(X, Y, X', Y') = -\sum_{x\in X(\Omega)=X'(\Omega)} \sum_{y\in Y(\Omega)=Y'(\Omega)} \mathbf{P}(X = x, Y = y) \ln \frac{\mathbf{P}(X' = x, Y' = y)}{\mathbf{P}(X = x, Y = y)}$$

en supposant que pour tout x, y : $\mathbf{P}(X = x, Y = y) \neq 0$ et $\mathbf{P}(X' = x, Y' = y) \neq 0$.
Dans cette partie, m et n sont des entiers naturels non nuls.

1. Propriétés
   (a) Montrer que
   $$K(X, Y, X', Y') = -\sum_{\substack{x \in X(\Omega) = X'(\Omega) \\ y \in Y(\Omega) = Y'(\Omega)}} \mathbf{P}(X = x, Y = y)\left(\ln \frac{\mathbf{P}(X' = x, Y' = y)}{\mathbf{P}(X = x, Y = y)} - \frac{\mathbf{P}(X' = x, Y' = y)}{\mathbf{P}(X = x, Y = y)} + 1\right)$$
   On peut avoir intérêt à noter $p_{x,y} = \mathbf{P}(X = x, Y = y)$ et $p'_{x,y} = \mathbf{P}(X' = x, Y' = y)$.
   (b) A l'aide de l'inégalité $(\mathcal{I}_1)$, établir que $K(X, Y, X', Y') \geqslant 0$ et que l'égalité a lieu si et seulement si les deux couples (X, Y) et (X', Y') ont la même loi conjointe.
   (c) On suppose que
   — X' et X suivent le même loi,
   — Y' et Y suivent le même loi,
   — X' et Y' sont indépendantes.
   Démontrer que $K(X, Y, X', Y') = \mathbf{S}(X) + \mathbf{S}(Y) - \mathbf{S}(X, Y)$.
   (d) Déduire de ce qui précède que $\mathbf{S}(X, Y) \leqslant \mathbf{S}(X) + \mathbf{S}(Y)$. Donner une condition nécessaire et suffisante pour que cette inégalité soit une égalité.

   On admet qu'on peut prolonger l'inégalité précédente au cas où certains des événements $(X = x, Y = y)$ sont négligeables.
2. Entropie conditionnelle.
   On définit l'entropie conditionnelle de Y sachant X par $\mathbf{S}_X(Y) = \mathbf{S}(X, Y) - \mathbf{S}(X)$.
   Elle mesure l'incertitude restant sur la valeur de Y lorsque la valeur de X est connue.
   (a) Montrer que $\mathbf{S}_X(Y) \leqslant \mathbf{S}(Y)$. Interpréter cette inégalité.
   (b) On considère $m + 1$ réels $a_0, a_1, \ldots a_m \in [0, 1]$.
      i. On suppose que pour tout $j \in [\![0, m]\!]$, $a_j \neq 0$. Démontrer alors que
      $$\forall\, j \in [\![0, m]\!], \qquad \ln(a_j) \leqslant \ln(\sum_{k=0}^{m} a_k)$$
      En déduire l'inégalité :
      $$\sum_{j=0}^{m} \varphi(a_j) \geqslant \varphi\left(\sum_{j=0}^{m} a_j\right) \qquad (\mathcal{I}_2)$$
      ii. L'inégalité $(\mathcal{I}_2)$ reste-t-elle vraie si $(a_0, a_1, \ldots a_m) \in [0, 1]^{m+1}$?
      iii. Montrer que l'inégalité $(\mathcal{I}_2)$ est une égalité si et seulement s'il existe au plus un indice $j \in [\![0, m]\!]$ pour lequel $a_j \neq 0$
   (c) Montrer que, pour tout $\forall\, x \in X(\Omega)$, $\sum_{y \in Y(\Omega)} \varphi\big(\mathbf{P}(X = x, Y = y)\big) \geqslant \varphi\big(\mathbf{P}(X = x)\big)$.
   En déduire que $\mathbf{S}_X(Y) \geqslant 0$