Devoir libre (n° 15)

Ce problème aborde la question de la mesure quantitative de l'information.

On cherche à modéliser l'information contenue dans les événements de probabilité non nulle d'un espace probabilisé. **L'information contenue dans un événement correspond intuitivement à l'effet de surprise provoqué par la réalisation de cet événement.**

On aborde ensuite la notion d'entropie d'une variable aléatoire discrète à valeurs entières, qui correspond intuitivement à la quantité moyenne d'information apportée par une réalisation de cette variable aléatoire.

On aborde ensuite l'entropie d'un couple de variables aléatoires et la notion d'entropie conditionnelle.

L'exemple traité dans la dernière partie établit le lien entre la quantité d'information contenue dans un message aléatoire et le nombre minimal de questions que le récepteur du message doit poser à son émetteur pour pouvoir identifier sans ambiguïté l'une des réalisations de ce message.

Questions préliminaires

1. Démontrer que, pour tout $x \in \mathbb{R}_+^*$, $\ln(x) \leqslant x - 1$ et que $\ln(x) = x - 1$ si et seulement si $x = 1$.

2. Montrer que la fonction g définie sur $[0,1]$ par $g(0) = 0$ et $\forall x \in]0,1]$, $g(x) = x\ln(x)$ est continue sur $[0,1]$ et dérivable sur $]0,1]$. Représenter graphiquement la fonction g.

Partie A – Mathématisation de l'effet de surprise

Soit (Ω,P) un espace probabilisé fini. On convient de modéliser la quantité d'information contenue dans les événements de probabilité non nulle par une fonction S définie, pour tout $A \in \mathcal{P}(\Omega)$ de probabilité non nulle, par $S(A) = f(P(A))$ où $f\colon]0,1] \to \mathbb{R}$ vérifie les contraintes suivantes :

(C1) $f(1) = 0$

(C2) f est décroissante sur $]0,1]$

(C3) $\forall (p,q) \in]0,1]^2 \quad f(pq) = f(p) + f(q)$

(C4) f est continue sur $]0,1]$

3. Quelle est la quantité d'information de l'événement certain ? Interpréter en terme d'effet de surprise.

4. Que peut-on dire de la quantité d'information contenue dans l'événement $A \cap B$ lorsque A et B sont indépendants ? Interpréter en terme d'effet de surprise.

5. Donner un exemple de fonction f vérifiant les quatre contraintes (C1), (C2), (C3) et (C4).

6. On se propose de déterminer l'ensemble des fonctions vérifiant ces quatre contraintes.
Soit f une telle fonction.

a) Soit $p \in]0,1]$. Établir, à l'aide d'un changement de variable, l'égalité

$$\frac{1}{p}\int_{p/2}^{p} f(t)\,\mathrm{d}t = \frac{1}{2}f(p) + \int_{1/2}^{1} f(u)\,\mathrm{d}u$$

b) En déduire que f est dérivable sur $]0,1]$.

c) Démontrer l'existence d'un réel a indépendant de p tel que $f'(p) = a/p$ pour $p \in \ ]0;1]$, puis préciser la valeur de a.

d) Déterminer l'ensemble des fonctions f vérifiant les quatre contraintes (C1), (C2), (C3) et (C4).

e) Montrer que parmi ces fonctions, il en existe une et une seule vérifiant en plus l'égalité $f(1/e) = 1$.

Cette fonction, notée h dans la suite du problème, correspond au choix d'une unité particulière (le *logon*) pour mesurer la quantité d'information.

Que vaut $\lim\limits_{p \to 0^+} h(p)$? Interpréter ce résultat.

f) On réalise l'expérience aléatoire consistant à effectuer deux lancers successifs d'un dé équilibré à six faces. On considère les événements suivants

— E : « le numéro sorti lors du premier lancer est pair » ;
— M : « le maximum des deux numéros sortis est égal à 4 » ;
— N : « la somme des deux numéros sortis est égale à 7 ».

Ordonner les quantités d'information contenues dans chacun de ces trois événements. Interpréter en terme d'effet de surprise.

Partie B – Entropie d'une variable aléatoire

Dans cette partie, toutes les variables aléatoires considérées sont définies sur un même univers fini Ω et prennent leurs valeurs dans $[\![0;n]\!]$.

Si X est une telle variable, on note $p_k = P(X = k)$. On définit l'entropie de X par

$$H(X) = -\sum_{k=0}^{n} p_k \ln(p_k)$$

en convenant que $p_k \ln(p_k)$ vaut 0 lorsque $p_k = 0$.

7. Interpréter $H(X)$ en terme de quantité d'information.

8. Montrer que $H(X) \geqslant 0$, puis montrer que $H(X) = 0$ si et seulement si X est une variable aléatoire certaine, c'est-à-dire s'il existe $i \in [\![0;n]\!]$ tel que $P(X = i) = 1$.

9. Soit X_0 une variable aléatoire suivant la loi uniforme sur $[\![0;n]\!]$. Calculer $H(X_0)$.

10. En utilisant une des questions préliminaires, démontrer que

$$\forall k \in [\![0;n]\!], \quad -p_k \ln(p_k) + p_k \ln\left(\frac{1}{n+1}\right) \leqslant \frac{1}{n+1} - p_k$$

11. En déduire que $H(X) \leqslant H(X_0)$ avec égalité si et seulement si X suit la même loi que X_0. Interpréter ce résultat en terme de quantité moyenne d'information.

Partie C – Entropie d'un couple de variables aléatoires

Dans cette partie, m et n sont des entiers naturels non nuls. (X,Y) et (X',Y') sont deux couples de variables aléatoires discrètes. X et X' sont à valeurs dans $[\![0;n]\!]$, Y et Y' sont à valeurs dans $[\![0;m]\!]$. Pour $i \in [\![0;n]\!]$ et $j \in [\![0;m]\!]$, on note $p_i = P(X = i)$, $q_j = P(Y = j)$, $\lambda_{ij} = P([X = i] \cap [Y = j])$ et $\lambda'_{ij} = P([X' = i] \cap [Y' = j])$.

On suppose que, pour tout $(i,j) \in [\![0;n]\!] \times [\![0;m]\!]$, $\lambda_{ij} \neq 0$ et $\lambda'_{ij} \neq 0$.

On définit l'entropie du couple (X,Y) par

$$H(X,Y) = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij} \ln(\lambda_{ij})$$

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

$$K(X,Y,X',Y') = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij} \ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right)$$

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

a) Rappeler les valeurs de $\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij}$ et de $\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda'_{ij}$ et en déduire que

$$K(X,Y,X',Y') = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij}\left(\ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right) - \frac{\lambda'_{ij}}{\lambda_{ij}} + 1\right)$$

b) 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 les deux variables X' et Y' sont indépendantes, que X' suit la même loi que X et que Y' suit la même loi que Y.
Démontrer que $K(X,Y,X',Y') = H(X) + H(Y) - H(X,Y)$. Déduire de ce qui précède que

$$H(X,Y) \leqslant H(X) + H(Y)$$

Donner une condition nécessaire et suffisante pour que cette inégalité soit une égalité.
Remarque — Cette inégalité a été obtenue en supposant, pour tout $(i,j) \in [\![0;n]\!] \times [\![0;m]\!]$, $\lambda_{ij} \neq 0$ et $\lambda'_{ij} \neq 0$. On admet qu'elle reste vraie même en dehors de cette condition.

14. *Entropie conditionnelle.*
On définit l'entropie conditionnelle de Y sachant X par $H_X(Y) = H(X,Y) - H(X)$.
Elle mesure l'incertitude restant sur la valeur de Y lorsque la valeur de X est connue.

a) Montrer que $H_X(Y) \leqslant H(Y)$. Interpréter cette inégalité.

b) On considère $m+1$ réels a_0, a_1, …, a_m appartenant à $]0;1]$.
Démontrer que, pour tout $j \in [\![0;m]\!]$, $\ln(a_j) \leqslant \ln(a_0 + a_1 + \cdots + a_m)$. En déduire l'inégalité

$$\sum_{j=0}^{m} g(a_j) \leqslant g\left(\sum_{j=0}^{m} a_j\right)$$

(La fonction g a été définie dans les question préliminaires.)
Cette inégalité reste-t-elle vraie si $(a_0,a_1,\ldots,a_m) \in [0,1]^{m+1}$?

c) Montrer que l'inégalité précédente est une égalité si et seulement s'il existe au plus un indice $j \in [\![0;m]\!]$ pour lequel $a_j \neq 0$.

d) Montrer que, pour tout $i \in [\![0;n]\!]$, $\displaystyle\sum_{j=0}^{m} g(\lambda_{ij}) \leqslant g(p_i)$. En déduire que $H_X(Y) \geqslant 0$.

Partie D – Une application

Un jeu oppose deux joueurs A et B. Une urne contient 2018 boules indiscernables au toucher et numérotées de 0 à 2017. Le joueur A tire une boule au hasard dans l'urne. On note Y la variable aléatoire égale au numéro de la boule tirée par le joueur A. Le joueur B pose alors au joueur A une série de N questions amenant la réponse « oui » ou « non » lui permettant de déterminer sans ambiguïté la valeur prise par Y. Le but de cette partie est de trouver la valeur minimale du nombre N de questions nécessaires pour que B soit certain de la valeur de Y.

1. Déterminer l'entropie de Y.

2. *Une stratégie possible : la dichotomie.*

La première question posée par B est « Y est-il compris entre 1009 et 2017 ? » La réponse fournie par A lui permet de positionner Y par rapport 1009. Si la réponse de A est « oui », B posera $X_1 = 1$ et cherchera ensuite à positionner Y par rapport à 1513. Si la réponse à la première question est « non », B posera $X_1 = 0$ et cherchera à positionner Y par rapport à 504. Il continue selon le même procédé.

Expliquer pourquoi le joueur B finira par trouver la valeur de Y. Au bout de combien de questions ?

3. On se propose de donner une minoration de N, quelque soit la stratégie de B, en utilisant l'entropie. Les variables aléatoires $(X_1,X_2,\dots,X_N)$ sont renseignées par les réponses données par A à la première, la deuxième, …, la N-ième question, elles prennent leurs valeurs dans $\{0,1\}$.

On pose $X = \displaystyle\sum_{i=1}^{N} X_{N+1-i}2^{i-1}$

Montrer que X est une variable aléatoire qui prend ses valeurs dans $[\![0; 2^N - 1]\!]$.

On ne cherchera pas la loi de X.

4. Démontrer que $H(X) \leqslant N\ln(2)$.

5. Expliquer en langage courant pourquoi $H_Y(X) = 0$. En déduire que $H(X,Y) = H(Y)$.

6. Montrer que $H_X(Y) \geqslant \ln(2018) - N\ln(2)$.

7. De combien de questions au minimum le joueur B a-t-il besoin pour être certain de pouvoir trouver la valeur de Y ?