Banque << Agro -Véto >>
A - 0508

MATHÉMATIQUES

ÉPREUVE B

Durée : 3 heures 30 minutes

L'usage d'une calculatrice est interdit pour cette épreuve. Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il le signale sur sa copie et poursuit sa composition en expliquant les raisons des initiatives qu'il a été amené à prendre.

L'objet de ce problème est l'étude de quelques propriétés élémentaires de la fonction d'entropie qui mesure l'incertitude sur la valeur prise par une variable aléatoire donnée.

NOTATIONS : Dans tout le problème n désigne un entier naturel.
On note $I_n = \{k \in \mathbb{N}, 0 \leqslant k \leqslant n\}$ l'ensemble des entiers naturels compris entre 0 et n.
Si f et g sont deux fonctions, on note, lorsque cela a un sens, $f \circ g$ leur composée.

PRÉLIMINAIRE.

1. On note h l'application de $\mathbb{R}^+$ dans $\mathbb{R}$ définie par $h(0) = 0$ et pour tout $x > 0$, $h(x) = x\ln(x)$. Montrer que h est continue sur $\mathbb{R}^+$.

2. Montrer que, pour tout $x > 0$, $\ln(x) \leqslant x - 1$ et que $\ln(x) = x - 1$ si et seulement si $x = 1$.

Dans le problème, on distingue deux cas :
1) Si, n étant un entier, X est une variable aléatoire discrète à valeurs dans I_n, de loi $p_i = P[X = i]$ pour tout entier $i \in I_n$, on note

$$H(X) = -\sum_{i=0}^{n} h(p_i)$$

l'entropie de la variable X.

2) Si X est une variable aléatoire à valeurs réelles de densité Φ, on dit que X possède une entropie si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} |h \circ \Phi(x)|dx$ est convergente. et on note alors

$$H(X) = -\int_{-\infty}^{+\infty} h \circ \Phi(x)dx$$

l'entropie de la variable X.

PARTIE I. EXEMPLES.

1. **Une variable discrète**.

Dans cette question, X est une variable aléatoire suivant la loi uniforme sur I_n. Calculer $H(X)$.

2. **Trois variables continues**.

2.1 Dans cette question, X est une variable aléatoire suivant la loi uniforme sur l'intervalle $[a, b]$ pour $a < b$.
Montrer que X possède une entropie et vérifier que $H(X) = \ln(b - a)$.

2.2. Dans cette question, X suit la loi exponentielle de paramètre $\lambda > 0$.
Montrer que X possède une entropie et vérifier que $H(X) = 1 - \ln(\lambda)$.

2.3. Dans cette question, X suit la loi normale de paramètres m et σ^2 où $m \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$, c'est à dire qu'une densité de X est $\Phi(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$.

Montrer que X possède une entropie et que $H(X) = \dfrac{1 + \ln(2\pi\sigma^2)}{2}$.

PARTIE II. PROPRIÉTÉS DE L'ENTROPIE

1. **Lois discrètes**

Dans cette question X_0 est une variable aléatoire de loi uniforme sur I_n et X désigne une variable aléatoire discrète à valeurs dans I_n.

1.1. Montrer que $H(X) \geqslant 0$ puis que $H(X) = 0$ si et seulement si X est une variable aléatoire certaine.

1.2. En utilisant le préliminaire, montrer que pour tout $i \in I_n$, $\quad -h(p_i) + p_i \ln(\frac{1}{n+1}) \leqslant \frac{1}{n+1} - p_i$.

1.3. En déduire que $H(X) \leqslant H(X_0)$, puis que $H(X) = H(X_0)$ si et seulement si X suit la loi uniforme.

1.4. Donner une interprétation de ce résultat.

2. **Lois continues**

Dans cette question, X_0 désigne une variable aléatoire de loi normale de paramètres m et σ^2 où $m \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$; on note Φ_0 sa densité.
X désigne une variable aléatoire de densité continue Φ, d'espérance m et de variance σ^2.
On suppose que X possède une entropie.

2.1. Calculer $I = \int_{-\infty}^{+\infty} \Phi(x) \ln(\Phi_0(x)) dx$ et vérifier que $H(X_0) = -I$.

2.2. En utilisant le préliminaire, montrer que pour tout $x \in \mathbb{R}$,

$$-h \circ \Phi(x) + \Phi(x) \ln[\Phi_0(x)] \leqslant \Phi_0(x) - \Phi(x).$$

2.3. En déduire $H(X) \leqslant H(X_0)$.

2.4. On suppose que $H(X_0) = H(X)$.

  2.4.1. Montrer que pour tout $x \in \mathbb{R}$, $\quad -h \circ \Phi(x) + \Phi(x) \ln[\Phi_0(x)] = \Phi_0(x) - \Phi(x)$.

  2.4.2. En déduire que X suit la loi normale de paramètres m et σ^2.

2.5. Donner une interprétation de ce résultat.

PARTIE III. ENTROPIE D'UN COUPLE

Dans cette partie n et m sont des entiers naturels non nuls, (X, Y) et (X', Y') sont deux couples de variables aléatoires discrètes.
Les variables aléatoires X et X' (resp. Y et Y') sont à valeurs dans I_n (resp. I_m).
Pour tout $i \in I_n$, on note p_i la probabilité $p_i = P[X = i]$, pour tout $j \in I_m$, on note q_j la probabilité $q_j = P[Y = j]$.
Enfin, pour tout couple $(i, j) \in I_n \times I_m$, on note $\lambda_{i,j}$ la probabilité $\lambda_{i,j} = P[X = i; Y = j]$ et $\mu_{i,j}$ la probabilité $\mu_{i,j} = P[X' = i; Y' = j]$.

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

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

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

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

$$K(X,Y,X',Y') = -\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{i,j}\ln(\frac{\mu_{i,j}}{\lambda_{i,j}}).$$

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

1.1. Rappeler les valeurs de $\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{m} \lambda_{i,j}$ et de $\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{m} \mu_{i,j}$.

1.2. Montrer que $K(X,Y,X',Y') = -\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{m} \lambda_{i,j}(\ln(\frac{\mu_{i,j}}{\lambda_{i,j}}) - \frac{\mu_{i,j}}{\lambda_{i,j}} + 1)$.

1.3. Montrer que $K(X,Y,X',Y') \geqslant 0$.

1.4. Montrer que $K(X,Y,X',Y') = 0$ si et seulement si les deux couples (X,Y) et (X',Y') ont même loi conjointe.

1.5. Dans cette question, on suppose que les deux variables aléatoires X' et Y' sont indépendantes, respectivement de même loi que X et Y. Pour tout i de I_n, on notera $p_i = P[X=i] = P[X'=i]$, respectivement pour tout j de I_m, $q_j = P[Y=j] = P[Y'=j]$.

1.5.1. Montrer que $K(X,Y,X',Y') = H(X) + H(Y) - H(X,Y)$.

1.5.2. Déduire de ce qui précède que $H(X,Y) \leqslant H(X) + H(Y)$ (inégalité (∇))

1.5.3. Donner une condition nécessaire et suffisante pour que cette inégalité soit une égalité.

REMARQUE : *l'inégalité* (∇) *a été obtenue en supposant, pour tout* (i,j) *de* $I_n \times I_m$, $\lambda_{i,j} \neq 0$ *et* $\mu_{i,j} \neq 0$. *On admet qu'elle demeure vraie sans cette condition.*

2. **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, la valeur de X étant connue.

2.1. Montrer que $H(Y|X) \leqslant H(Y)$.

2.2. On considère $m+1$ réels $a_0, a_1, \ldots, a_m$ compris entre 0 et 1.

2.2.1. Dans cette question, on suppose $(a_0, a_1, \ldots, a_m) \in]0,1]^{m+1}$.
Montrer que pour tout i de I_m, $\ln(a_i) \leqslant \ln(a_0 + a_1 + \cdots + a_m)$.
En déduire que $\sum\limits_{i=0}^{m} h(a_i) \leqslant h\left(\sum\limits_{i=0}^{m} a_i\right)$. (inégalité (*))

2.2.2. L'inégalité (*) demeure-t-elle si $(a_0, a_1, \ldots a_m) \in [0,1]^{m+1}$?

2.2.3. Montrer que l'inégalité (*) est une égalité si et seulement s'il existe au plus un indice $i \in I_m$ pour lequel $a_i \neq 0$.

2.3. Montrer que pour tout $i \in I_n$, $\sum\limits_{j=0}^{m} h(\lambda_{i,j}) \leqslant h(p_i)$. En déduire que $H(Y|X) \geqslant 0$.

2.4. Dans cette question, on suppose $H(Y|X) = 0$.
Montrer que, pour tout $i \in I_n$ tel que $p_i > 0$, il existe un unique $j \in I_m$ tel que $\lambda_{i,j} > 0$. On notera $j = \alpha(i)$.
Montrer que $Y = \alpha(X)$. Comment interpréter ce résultat ?

On admettra pour la suite la réciproque de la propriété démontrée en 2.4 ci-dessus.

PARTIE IV. UNE APPLICATION.

Un jeu oppose deux joueurs. Le joueur A tire une boule au hasard dans une urne contenant 2009 boules numérotées de 0 à 2008. On note Y la variable aléatoire égale au numéro de la boule tirée.
Dans le but de déterminer Y, le joueur B pose une série de N question(s) où N est un entier naturel non nul fixé à l'avance.
Ces questions sont impérativement de la forme : « le nombre choisi appartient-il à E ?» où E est un sous-ensemble de I_{2008} pouvant varier à chaque question.
Le joueur A répond par l'affirmative si $Y \in E$, par la négative sinon.
Au bout des N question(s), B donne sa réponse.
En cours de jeu, on note $X_i = 1$ si la réponse de A est affirmative à la $i$ème question de B et $X_i = 0$ sinon.

1. On pose $X = \sum_{i=1}^{N} X_i 2^{i-1}$. On admet que X est une variable aléatoire. Montrer qu'elle est à valeurs dans $I_{(2^N-1)}$ (On ne cherchera pas à déterminer sa loi.)

2. Déterminer l'entropie $H(Y)$

3. Montrer que $H(X) \leqslant N\ln(2)$ (On utilisera l'inégalité vue à la question II.1.3.)

4. Expliquer, en langage courant, pourquoi on a $H(X|Y) = 0$. En déduire que $H(X,Y) = H(Y)$.

5. Montrer que $H(Y|X) \geqslant \ln(2009) - N\ln(2)$.

6. Le joueur B prétend pouvoir trouver la réponse à coup sûr en 10 questions (ou moins). Qu'en pensez-vous ? En combien de questions pouvez-vous espérer donner à coup sûr la valeur de Y ?

FIN DE L'ÉPREUVE