COMPOSITION DE MATHÉMATIQUES

(Sujet commun ENS : ULM, LYON et CACHAN) (C)

DURÉE : 3 heures

Bien que traitant d'un thème commun, les trois parties du problème peuvent être abordées indépendamment, mises à part quelques notations introduites dans le préliminaire de la première partie. Une lecture attentive des préliminaires de chaque partie est vivement conseillée.

Dans tout le problème ln désigne le logarithme népérien, exp l'exponentielle de base e et ψ la fonction définie sur $\mathbb{R}^+$ par :

$$\psi(0) = 0 \quad \text{et} \quad \psi(x) = x \ln(x) \quad \text{si} \quad x > 0.$$

PREMIÈRE PARTIE

Soient $(\Omega, \mathcal{F}, P)$ un espace de probabilité, X et Y deux variables aléatoires définies sur Ω à valeurs dans un même ensemble fini E, dont le nombre d'éléments est noté $|E|$. On dit que Y est vraisemblable pour X s'il existe une fonction L de E dans $\mathbb{R}^+$, appelée vraisemblance de Y pour X telle que :

$$\text{pour } a \in E, \quad P(Y = a) = L(a)\, P(X = a); \quad L(a) = 0 \quad \text{si} \quad P(X = a) = 0.$$

On remarquera que Y est vraisemblable pour X si et seulement si :

$$\text{pour } a \in E, \quad [P(X = a) = 0] \Rightarrow [P(Y = a) = 0].$$

et que la vraisemblance de Y pour X est alors unique (en particulier si $P(X = a) \neq 0$ pour tout $a \in E$, toute variable Y est vraisemblable pour X).

Si Y est vraisemblable pour X de vraisemblance L, on appelle entropie relative de Y pour X la quantité :

$$K(Y, X) = \sum_{a \in E} \psi(L(a))\, P(X = a) = E[\psi(L(X))].$$

(E [Z] désigne l'espérance mathématique de la variable aléatoire Z).

On appelle entropie de X la quantité :

$$H(X) = -\sum_{a \in E} \psi(P(X = a)).$$

I.1.*a.* Montrer que pour x et y dans $\mathbb{R}^+$ on a $\ln(x) - \ln(y) \leqslant \dfrac{x}{y} - 1$.

b. En déduire que pour tout $x \geqslant 0$, $\psi(x) + 1 - x \geqslant 0$.

I.2.*a.* Calculer H (X) lorsque X est constante, puis lorsque X suit la loi uniforme sur E.

b. Quel est le signe de H (X) ?

I.3.*a.* On suppose Y vraisemblable pour X de vraisemblance L. Montrer que $K(Y, X) \geqslant 0$ (on pourra utiliser 1.*b.*). Quand y a-t-il égalité ?

b. On suppose que X suit la loi uniforme sur E. Montrer que toute variable Y est vraisemblable pour X et que $K(Y, X) = H(X) - H(Y)$.

En déduire une majoration de H (Y).

I.4. On suppose Y vraisemblable pour X de vraisemblance L.

a. Soit u une fonction de E dans $\mathbb{R}_+^*$. Montrer que :

$$E\left[(L(X)\ln(u(X))) - \psi(L(X))\right] \leqslant E\left[u(X) - 1\right]$$

(on pourra utiliser 1.*a.*). Pour quelle(s) u y a-t-il égalité ?

b. En déduire que pour toute fonction v de E dans $\mathbb{R}$, on a :

$$E\left[v(Y)\right] - E\left[\exp(v(X))\right] + 1 \leqslant K(Y,X).$$

Pour quelle(s) v y a-t-il égalité ?

c. On note $h(v)$ la quantité $E\left[v(Y)\right] - E\left[\exp(v(X))\right]$. Pour $\lambda \in \mathbb{R}$ on note $\overline{\lambda}$ la fonction constante sur E égale à λ. On pose $g_v(\lambda) = h(\overline{\lambda} + v)$.

Montrer que g_v admet un unique maximum. En déduire :

$$E\left[v(Y)\right] - \ln\left(E[\exp(v(X))]\right) \leqslant K(Y,X).$$

Pour quelle(s) v y a-t-il égalité ?

I.5. Dans cette question $E = E_1 \times E_2$, $X = (X_1, X_2)$, $Y = (Y_1, Y_2)$ où les X_i, Y_i sont à valeurs dans E_i $(i = 1,2)$.

a. On suppose ici $E_1 = \{1,2\}$, $E_2 = \{1,2,3\}$ et on donne la loi de X suivante :

$$P(X = (1,j)) = \frac{1}{6} \quad \text{pour} \quad j = 1,2,3;$$

$$P(X = (2,3)) = \frac{1}{2};$$

$$P(X = (2,j)) = 0 \quad \text{pour} \quad j = 1,2.$$

Quelles sont les lois de X_1 et X_2 ? X_1 et X_2 sont-elles indépendantes ? Montrer qu'il n'existe pas de variable Y vraisemblable pour X telle que ses lois marginales (c'est-à-dire les lois de Y_1 et Y_2) soient les lois uniformes sur E_1 et E_2.

b. On suppose maintenant que X suit la loi uniforme sur E. Quelles sont les lois de X_1 et X_2 ? X_1 et X_2 sont-elles indépendantes ?

Montrer que toute Y est vraisemblable pour X, puis que :

$$K(Y,X) \geqslant K(Y_1,X_1) + K(Y_2,X_2)$$

(on pourra utiliser 4.*c.*) avec $v(a_1,a_2) = v_1(a_1) + v_2(a_2)$ pour un choix judicieux de v_1 et v_2).

Montrer qu'il y a égalité lorsque Y_1 et Y_2 sont indépendantes.

En déduire que $H(Y) \leqslant H(Y_1) + H(Y_2)$ avec égalité lorsque Y_1 et Y_2 sont indépendantes.

DEUXIÈME PARTIE

Dans cette partie $E = \mathbb{N}$, donc n'est plus fini. On suppose que X suit une loi de Poisson de paramètre λ ($\lambda > 0$).

Y est une variable aléatoire à valeurs dans $\{0, \ldots, n\}$ pour un $n \in \mathbb{N}$. Puisque $P(X = k) \neq 0$ pour tout $k \in \mathbb{N}$, Y est vraisemblable pour X (au sens de la première partie) et sa vraisemblance L vérifie $L(k) = 0$ si $k \geqslant n+1$. On peut donc encore définir l'entropie relative de Y pour X par :

$$K(Y, X) = \sum_{k=0}^{n} \psi(L(k))\, P(X = k) = E[\psi(L(X))].$$

II.1. Pour $n \in \mathbb{N}^*$ et $n \geqslant \lambda$, Y_n est une variable aléatoire qui suit la loi binomiale de paramètres n et $\frac{\lambda}{n}$. On note L_n sa vraisemblance pour X.

a. Montrer que pour $k \in \{0, \ldots, n-1\}$, $L_n(k+1) = L_n(k) \dfrac{n-k}{n-\lambda}$.

En déduire que :

$$K(Y_n, X) = \ln(L_n(0)) + \sum_{k=1}^{n} P(Y_n = k) \left(\sum_{i=0}^{k-1} \ln\left(\frac{1 - (i/n)}{1 - (\lambda/n)} \right) \right).$$

b. Déterminer $L_n(0)$. Montrer alors que :

$$K(Y_n, X) \leqslant \lambda + (n - \lambda) \ln\left(1 - \frac{\lambda}{n}\right)$$

(on pourra regarder le signe de $\sum_{i=0}^{k-1} \ln(1 - (i/n))$).

En déduire que $K(Y_n, X)$ tend vers 0 quand n tend vers l'infini.

II.2. On note $d(Y, X) = \sum_{k=0}^{\infty} |P(Y = k) - P(X = k)|$.

a. Vérifier que $d(Y, X)$ est bien défini et que $d(Y, X) = E[|L(X) - 1|]$.

b. Pour $x \geqslant 0$, on pose $\varphi(x) = (\psi(x) + 1 - x)(4 + 2x) - 3(x-1)^2$.

Calculer $\varphi'(x)$ (pour $x > 0$), $\varphi'(1)$, $\varphi(1)$. Étudier les variations de φ' (on pourra utiliser I.1.*b.*).

En déduire que $\varphi(x) \geqslant 0$. Quand y a-t-il égalité ?

c. On rappelle que si Z et Z' sont des variables aléatoires positives ou nulles, $E[\sqrt{ZZ'}] \leqslant \sqrt{E[Z]\, E[Z']}$.

Déduire de ce qui précède que :

$$d(Y, X) \leqslant \sqrt{2K(Y, X)}.$$

II.3. Montrer à l'aide de 1. et 2. que $d(Y_n, X)$ tend vers 0 quand n tend vers l'infini. Qu'en déduit-on pour $P(Y_n = k)$, k fixé ?

TROISIÈME PARTIE

Dans cette partie on considère des variables aléatoires réelles (X, Y, ...) dont la loi admet une densité ($f_X, f_Y, \ldots$) strictement positive.

La vraisemblance de Y pour X est alors définie pour $t \in \mathbb{R}$ par $L(t) = \dfrac{f_Y(t)}{f_X(t)}$, et l'entropie relative de Y pour X est définie par :

$$K(Y, X) = E[\psi(L(X))] = \int_{-\infty}^{+\infty} \psi(L(t))\, f_X(t)\, dt$$

lorsque cette intégrale est absolument convergente.

III.1. Pour $\theta > 0$, on considère des variables X_θ dont la loi a pour densité $f_\theta(t) = \theta \exp(-2\theta|t|)$.

a. Calculer $K(\theta, \theta') = K(X_\theta, X_{\theta'})$ pour θ et θ' strictement positifs.

b. Montrer que $\lim\limits_{\theta' \to \theta} \dfrac{K(\theta, \theta')}{(\theta - \theta')^2}$ existe et la calculer.

c. On considère la densité précédente comme une fonction de deux variables $f(\theta, t) = \theta \exp(-2\theta|t|)$, $\theta > 0$ et $t \in \mathbb{R}$.

Calculer $\dfrac{\partial f}{\partial \theta}(\theta, t)$, puis :

$$J(\theta) = \int_{-\infty}^{+\infty} \left(\frac{\partial f}{\partial \theta} / f\right)^2 (\theta, t)\, f(\theta, t)\, dt.$$

N.B. — Dans la formule précédente / désigne le quotient.

d. Comparer $J(\theta)$ et la limite obtenue au *b.*

III.2. Reprendre la question précédente pour des variables X_θ suivant une loi gaussienne, centrée, de variance $\theta > 0$. Le résultat du 1.*d.* subsiste-t-il ?