BANQUE COMMUNE D'ÉPREUVES

Code sujet : 280

Conception : HEC Paris

MATHÉMATIQUES

OPTION SCIENTIFIQUE

Mercredi 30 avril 2014, de 8 h. à 12 h.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à **encadrer** dans la mesure du possible les résultats de leurs calculs.
Ils ne doivent faire usage d'aucun document : **l'utilisation de toute calculatrice et de tout matériel électronique est interdite.** Seule l'utilisation d'une règle graduée est autorisée.
Si au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre

---

Dans ce problème, on s'intéresse à des opérations de transport dans des situations déterministes ou aléatoires, modélisées de manière discrète ou continue, dans le but de trouver un programme de transport optimal dont le coût serait le plus faible possible.

Les parties I, II et III sont largement indépendantes.

- Toutes les variables aléatoires considérées dans ce problème sont supposées définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$.
- Sous réserve d'existence, on note $E(Z)$ l'espérance d'une variable aléatoire Z.
- Pour tout entier N supérieur ou égal à 1, on note $\mathcal{E}_N$ l'ensemble des applications de $[\![1, N]\!]$ dans $[\![1, N]\!]$.

Préliminaire

1. Soit N un entier supérieur ou égal à 2.
   a) Quel est le nombre d'éléments de l'ensemble $\mathcal{E}_N$?
   b) Parmi les éléments de $\mathcal{E}_N$, quel est le nombre d'applications injectives et parmi celles-ci, combien sont strictement monotones ?
   (les réponses aux questions 1.a) et 1.b) seront données sans démonstration)

2. Soit p un réel vérifiant $0 < p < 1$.
   On considère une variable aléatoire X suivant la loi exponentielle de paramètre 1.
   Pour tout $\omega \in \Omega$, on pose : $Y(\omega) = \lfloor p\,X(\omega) \rfloor$, où $\lfloor\ \ \rfloor$ désigne la fonction partie entière.
   a) Vérifier que Y est une variable aléatoire discrète. Calculer pour tout $n \in \mathbb{N}$, la probabilité $P\big([Y = n]\big)$.
   b) Montrer que la variable aléatoire $Y + 1$ suit une loi géométrique dont on précisera le paramètre.
   c) Établir les inégalités strictes : $0 < E(Y) < p$.

Tournez la page S.V.P.

3.a) Pour tout couple $(r,s) \in \mathbb{N}^2$, montrer que l'intégrale $\displaystyle\int_0^1 x^r (\ln x)^s \,\mathrm{d}x$ est convergente.

(on pourra utiliser le changement de variable $u = -\ln x$ après avoir justifié précisément sa validité)

b) Établir pour tout couple $(r,s) \in \mathbb{N}^2$, l'égalité : $\displaystyle\int_0^1 x^r (\ln x)^s \,\mathrm{d}x = \frac{(-1)^s s!}{(r+1)^{s+1}}$.

Partie I. Transport dans une situation aléatoire

On dit que la loi d'une variable aléatoire Y est *accessible* depuis une variable aléatoire X, s'il existe une application $T : X(\Omega) \longrightarrow \mathbb{R}$ telle que la variable aléatoire $T(X)$ suit la même loi que Y.
L'application T est alors appelée une *fonction de transport* de la variable aléatoire X vers la loi de Y.
On associe à T un *coût de transport* $C(T)$ défini, sous réserve d'existence, par : $C(T) = E\Big(\big(X - T(X)\big)^2\Big)$.

Dans toute cette partie, X désigne une variable aléatoire vérifiant $X(\Omega) =]0,1[$ et suivant la loi uniforme sur $]0,1[$, c'est-à-dire admettant pour densité la fonction f_X définie par :

$$f_X(x) = \begin{cases} 1 & \text{si } x \in]0,1[\\ 0 & \text{sinon} \end{cases}.$$

4. Soit p un réel vérifiant $0 < p < 1$. Pour tout réel $a \in [0, 1-p]$, on note dans cette question, T_a la fonction définie sur $]0,1[$ par :

$$T_a(x) = \begin{cases} 1 & \text{si } x \in]a, a+p[\\ 0 & \text{sinon} \end{cases}.$$

a) Calculer la probabilité $P\big([T_a(X) = 1]\big)$ et en déduire que les fonctions T_a sont des fonctions de transport de X vers une même loi que l'on précisera.

b) Vérifier que le coût de transport $C(T_a)$ est égal à $\dfrac{1}{3} + p(1-p) - 2ap$.

c) En déduire la valeur de a qui minimise $C(T_a)$ et exprimer le coût minimal correspondant en fonction de p.

5. Soit T_1 et T_2 les applications définies sur $]0,1[$ par $T_1(x) = -\ln x$ et $T_2(x) = -\ln(1-x)$.

a) Vérifier que T_1 et T_2 sont des fonctions de transport de X vers une loi que l'on précisera.

b) En utilisant les résultats de la question 3, comparer les coûts de transport $C(T_1)$ et $C(T_2)$.

c) À l'aide de la question 2, montrer que toutes les lois géométriques sont accessibles depuis X.

6. Dans cette question, Y désigne une variable aléatoire admettant une densité f_Y continue et strictement positive sur $\mathbb{R}$.

a) Justifier que la fonction de répartition F_Y de Y réalise une bijection de $\mathbb{R}$ sur l'intervalle ouvert $]0,1[$.

b) On note F_Y^{-1} la bijection réciproque de F_Y.
Montrer que F_Y^{-1} est une fonction de transport de la variable aléatoire X vers la loi de Y.

7. *Cas particulier : on suppose que Y suit la loi normale centrée réduite.*

On note F_Y la fonction de répartition de Y et φ la densité continue sur $\mathbb{R}$ de Y.

a) Établir la convergence de l'intégrale $\displaystyle\int_{-\infty}^{+\infty} y\, F_Y(y)\, \varphi(y)\,\mathrm{d}y$.

À l'aide d'une intégration par parties, montrer que $\displaystyle\int_{-\infty}^{+\infty} y\, F_Y(y)\, \varphi(y)\,\mathrm{d}y = \frac{1}{2\sqrt{\pi}}$.

b) Montrer que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \left(y - F_Y(y)\right)^2 \varphi(y)\,\mathrm{d}y$ est convergente et la calculer.

c) En déduire que le coût de transport $C(F_Y^{-1})$ est égal à $\dfrac{4}{3} - \dfrac{1}{\sqrt{\pi}}$.

Partie II. Transport optimal dans une situation déterministe

Dans toute cette partie, N désigne un entier supérieur ou égal à 2.

On considère N réels $d_1, d_2, \ldots, d_N$ (appelés points de départ) et N réels $a_1, a_2, \ldots, a_N$ (appelés points d'arrivée) vérifiant $d_1 < d_2 < \cdots < d_N$ et $a_1 < a_2 < \cdots < a_N$.
On pose : $D = \{d_1, d_2, \ldots, d_N\}$ et $A = \{a_1, a_2, \ldots, a_N\}$.

8.a) Montrer que pour tout couple $(k, \ell) \in [\![1, N]\!]^2$, on a : $d_k a_k \geqslant d_k a_\ell + d_\ell a_k - d_\ell a_\ell$.

b) En déduire à l'aide d'une double sommation que pour tout N-uplet $(p_1, p_2, \ldots, p_N) \in \mathbb{R}_+^N$ tel que $\displaystyle\sum_{k=1}^{N} p_k = 1$, on a :

$$\sum_{k=1}^{N} p_k\, d_k\, a_k \geqslant \left(\sum_{k=1}^{N} p_k\, d_k\right) \times \left(\sum_{k=1}^{N} p_k\, a_k\right) \qquad (1)$$

9. Soit $t \in \mathcal{E}_N$. On réordonne la liste $\big(t(1), t(2), \ldots, t(N)\big)$ selon les valeurs croissantes et on note alors $\big(\widehat{t}(1), \widehat{t}(2), \ldots, \widehat{t}(N)\big)$ la liste ordonnée obtenue. On a donc : $\widehat{t}(1) \leqslant \widehat{t}(2) \leqslant \cdots \leqslant \widehat{t}(N)$.

a) Justifier pour tout $n \in [\![1, N]\!]$, l'inégalité : $\displaystyle\sum_{k=n}^{N} a_{\,t(k)} \leqslant \sum_{k=n}^{N} a_{\,\widehat{t}(k)}$.

b) On pose $d_0 = 0$. Justifier l'égalité : $\displaystyle\sum_{n=1}^{N} d_n\, a_{\,t(n)} = \sum_{n=1}^{N} \left((d_n - d_{n-1}) \sum_{k=n}^{N} a_{\,t(k)}\right)$.

c) Établir l'inégalité : $\displaystyle\sum_{n=1}^{N} d_n\, a_{\,t(n)} \leqslant \sum_{n=1}^{N} d_n\, a_{\,\widehat{t}(n)}$. $\qquad (2)$

On appelle *programme de transport*, toute bijection T de D sur A, et *coût* d'un programme de transport T, la somme $c(T)$ définie par : $c(T) = \displaystyle\sum_{k=1}^{N} \big(d_k - T(d_k)\big)^2$.

10. Soit $\widehat{T}$ le programme de transport défini par : pour tout $k \in [\![1, N]\!]$, $\widehat{T}(d_k) = a_k$.

Déduire des questions précédentes que le programme $\widehat{T}$ est optimal, c'est-à-dire que pour tout programme de transport T, on a : $c(T) \geqslant c(\widehat{T})$.

11. *Interprétation probabiliste des inégalités (1) et (2).*

Soit h une application croissante de $\mathbb{R}$ dans $\mathbb{R}$.

a) En utilisant l'inégalité (1), établir pour toute variable aléatoire discrète X ne prenant qu'un nombre fini de valeurs, l'inégalité : $E\big(X\,h(X)\big) \geqslant E(X)\,E\big(h(X)\big)$.

b) Que peut-on en déduire pour le coefficient de corrélation linéaire de X et $h(X)$ lorsque les variances de X et $h(X)$ sont strictement positives ?

c) En utilisant l'inégalité (2), montrer que si X est une variable aléatoire discrète suivant la loi uniforme sur $[\![1, N]\!]$ et t un élément de $\mathcal{E}_N$, on a : $E\big(h(X)\,t(X)\big) \leqslant E\big(h(X)\,\widehat{t}(X)\big)$.

Tournez la page S.V.P.

Partie III. Transport optimal dans une situation aléatoire

Les définitions de fonction de transport et de coût de transport sont identiques à celles données dans le préambule de la partie I.

Dans toute cette partie, U désigne une variable aléatoire vérifiant $U(\Omega) = [0, 1]$ et suivant la loi uniforme sur le segment $[0, 1]$.

Soit Y une variable aléatoire admettant une densité f_Y nulle hors d'un segment $[\alpha, \beta]$ $(\alpha < \beta)$ et dont la restriction à ce segment est continue et strictement positive. On note F_Y la fonction de répartition de Y.

On suppose l'existence d'une fonction g de classe C^1 sur $[0, 1]$, à valeurs dans $[\alpha, \beta]$, telle que la variable aléatoire $Z = g(U)$ suit la même loi que Y.

12. Pour tout entier $N \geqslant 1$, on pose pour tout $\omega \in \Omega$:

$$X_N(\omega) = \begin{cases} \lfloor 1 + N\,U(\omega) \rfloor & \text{si } 0 \leqslant U(\omega) < 1 \\ N & \text{si } U(\omega) = 1 \end{cases} \qquad \text{et} \qquad Y_N(\omega) = g\left(\frac{X_N(\omega)}{N}\right).$$

a) Trouver la loi de la variable aléatoire X_N.

b) Établir l'existence d'une constante $\lambda > 0$, indépendante de N telle que : $\forall \omega \in \Omega$, $|Z(\omega) - Y_N(\omega)| \leqslant \dfrac{\lambda}{N}$.

c) Montrer que pour tout réel y, on a : $F_Y\left(y - \dfrac{\lambda}{N}\right) \leqslant P([Y_N < y])$.

13. Pour tout $k \in [\![1, N]\!]$, on pose : $t_N(k) = g\left(\dfrac{k}{N}\right)$. On définit alors $\widehat{t}_N$ à partir de t_N, comme $\widehat{t}$ à partir de t dans la question 9.

a) Établir pour tout $k \in [\![1, N]\!]$, les inégalités : $F_Y\left(\widehat{t}_N(k) - \dfrac{\lambda}{N}\right) \leqslant P([Y_N < \widehat{t}_N(k)]) < \dfrac{k}{N}$.

b) On note F_Y^{-1} la fonction réciproque de la restriction à $[\alpha, \beta]$ de la fonction F_Y.

Montrer que pour tout entier $N \geqslant 1$, on a : $\displaystyle \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N}\, g\left(\frac{k}{N}\right) \leqslant \frac{1}{N}\sum_{k=1}^{N} \frac{k}{N}\left(F_Y^{-1}\left(\frac{k}{N}\right) + \frac{\lambda}{N}\right)$.

c) En déduire l'inégalité : $E\big(U\,g(U)\big) \leqslant E\big(U\,F_Y^{-1}(U)\big)$.

14.a) Parmi les fonctions de transport de classe C^1 de U vers la loi de Y, trouver une fonction de transport T^* de coût minimal.

b) On suppose que $Y = |4\,U - 2|$. Déterminer T^* et $C(T^*)$.

IMPRIMERIE NATIONALE – 14 1217 – D'après documents fournis

ANNALES DE MATHEMATIQUES 2014

HEC 2014 VOIE S

CORRIGE

Préliminaires

1–a)

$$\text{Card}(\mathcal{E}_N) = N^N$$

1–b)

Puisque $\mathcal{E}_N$ est est un ensemble fini, pour toute application appartenant à $\mathcal{E}_N$, il y a équivalence entre injection et bijection.

Le nombre d'applications injectives de $\mathcal{E}_N$ dans lui-même est $N!$.

Parmi celles-ci, il y en exactement deux qui sont strictement monotones : l'identité et l'application $(1, 2, \ldots, k, \ldots, N-1, N) \longrightarrow (N, N-1, \ldots N+1-k, \ldots, 2, 1)$ strictement décroissante.

2–a)

$X(\Omega) = \mathbb{R}_+$, donc $(pX)(\Omega) = \mathbb{R}_+$ et par suite $\lfloor (pX) \rfloor = \mathbb{N}$. La variable Y est une variable discrète.

$$\begin{aligned}
\forall n \in \mathbb{N},\ (Y = n) &= (n \leq pX < n+1) = \left(\frac{n}{p} \leq X < \frac{n+1}{p}\right) \quad \text{car } p > 0 \\
P(Y = n) &= \int_{\frac{n}{p}}^{\frac{n+1}{p}} e^{-t} dt = \Big[-e^{-t} \Big]_{\frac{n}{p}}^{\frac{n+1}{p}} \\
&= \exp(-\frac{n}{p}) - \exp(-\frac{n+1}{p}) \\
&= \exp(-\frac{n}{p})\Big(1 - \exp(-\frac{1}{p})\Big)
\end{aligned}$$

$$\forall n \in \mathbb{N},\ P(Y = n) = (\exp(-\frac{1}{p}))^n (1 - \exp(-\frac{1}{p}))$$

On peut remarquer que la condition $p < 1$ n'est pas intervenue. Nous utiliserons cette remarque dans la question II–5–c).

2–b)

$(Y+1)(\Omega) = \mathbb{N}^*$.

$\forall n \in \mathbb{N}^*,\ P(Y+1 = n) = P(Y = n-1) = (\exp(-\frac{1}{p}))^{n-1}(1 - \exp(-\frac{1}{p}))$.

$p > 0 \Longrightarrow 0 < 1 - \exp(-\frac{1}{p}) < 1$. Donc

$$\boxed{\text{La variable } Y+1 \text{ suit la loi géométrique de paramètre } 1 - \exp(-\frac{1}{p})}$$

2–c)

D'après le cours, $E(Y+1) = \dfrac{1}{1-\exp(-\frac{1}{p})}$, donc $E(Y) = \dfrac{1}{1-\exp(-\frac{1}{p})} - 1$,

$$\boxed{E(Y) = \frac{\exp(-\frac{1}{p})}{1-\exp(-\frac{1}{p})}}$$

$p > 0 \Longrightarrow \exp(-\frac{1}{p}) < 1$, donc $\dfrac{\exp(-\frac{1}{p})}{1-\exp(-\frac{1}{p})} > 0 \Longleftrightarrow E(Y) > 0$.

$$\begin{aligned} E(Y) < p \quad &\Longleftrightarrow \frac{\exp(-\frac{1}{p})}{1-\exp(-\frac{1}{p})} < p \\ &\Longleftrightarrow \exp(-\tfrac{1}{p}) < p(1-\exp(-\tfrac{1}{p})) \quad \text{car } 1-\exp(-\tfrac{1}{p}) > 0 \\ &\Longleftrightarrow \exp(-\tfrac{1}{p})(1+p) < p \\ &\Longleftrightarrow \exp(\tfrac{1}{p}) > \frac{p+1}{p} = 1 + \frac{1}{p}. \end{aligned}$$

On sait, par convexité de la fonction exponentielle, que $\forall x \in \mathbb{R},\ e^x \geq 1+x$ avec égalité si et seulement si $x = 0$, donc $\forall p \in]0,1[,\ 1+\frac{1}{p} < \exp(\frac{1}{p})$, ce qui équivaut à $E(Y) < p$. Finalement,

$$\boxed{0 < E(Y) < p}$$

3–a)

Soit $I(r,s) = \displaystyle\int_0^1 x^r(\ln x)^s dx$. La fonction $x \mapsto x^r(\ln x)^s$ est continue sur $]0,1]$. L'intégrale est impropre uniquement en 0.

- Si $r \geq 1$, $\displaystyle\lim_{x\to 0^+} x^r(\ln x)^s = 0$ par croissances comparées. L'intégrale est faussement impropre, donc convergente.
- Si $r = 0$, $\displaystyle\lim_{x\to 0^+} |\sqrt{x}(\ln x)^s| = 0$ par croissances comparées ; cela veut dire

$|(\ln x)^s| \underset{(0^+)}{=} \circ\Big(\dfrac{1}{\sqrt{x}}\Big)$. D'après le critère de Riemann, l'intégrale $\displaystyle\int_0^1 \frac{1}{\sqrt{x}}dx$ est convergente car $\sqrt{x} = x^{\frac{1}{2}}$ avec $\frac{1}{2} < 1$. Par comparaison des fonctions continues, positives pour les intégrales impropres, on conclut que l'intégrale $\displaystyle\int_0^1 (\ln x)^s dx$ est absolument convergente, donc convergente.

$$\boxed{\forall (r,s) \in \mathbb{N}^2,\ \ \text{l'intégrale } I(r,s) = \int_0^1 x^r(\ln x)^s dx \text{ est convergente}}$$

Effectuons le changement de variable $u = -\ln x$ qui est C^1 bijectif sur $]0,1]$, donc légitime.

$x = e^{-u} \Longrightarrow dx = -e^{-u}du$.

$$\begin{aligned} I(r,s) \quad &= \int_{+\infty}^0 (e^{-u})^r(-u)^s(-e^{-u})du \\ &= \int_0^{+\infty} (-1)^s u^s e^{-(r+1)u}du = (-1)^s \int_0^{+\infty} u^s e^{-(r+1)u}du \end{aligned}$$

3–b)

Posons $J(r,s) = \displaystyle\int_0^{+\infty} u^s e^{-(r+1)u}du$ et effectuons le changement de variable affine, donc licite, $t = (r+1)u$; $dt = (r+1)du$.