HEC Paris - ESSEC BS 2019 - voie E

Exercice

1. Dans cette question, on considère les matrices $C = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \in \mathcal{M}_{3,1}(\mathbf{R})$, $L = [1, 2, -1] \in \mathcal{M}_{1,3}(\mathbf{R})$ et le produit matriciel $M = CL$.
   (a) i. Calculer M et M^2.
   ii. Déterminer le rang de M.
   iii. La matrice M est-elle diagonalisable ?
   (b) i. Soit $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -2 & 1 \end{bmatrix}$. Justifier que P est inversible et calculer le produit $P \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$.
   ii. Trouver une matrice inversible Q dont la transposée ${}^tQ$ vérifie : ${}^tQ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.
   iii. Pour une telle matrice Q, calculer le produit PMQ.

2. La fonction *Scilab* suivante permet de multiplier la i-ème ligne L_i d'une matrice A par un réel sans modifier ses autres lignes, c'est-à-dire de lui appliquer l'opération élémentaire $L_i \leftarrow aL_i$ (où $a \neq 0$).

```
function B=multlig(a,i,A)
    [n,p] = size(A)
    B = A
    for j=1:p
        B(i,j)=a*B(i,j)
    end
endfunction
```

   (a) Donner le code *Scilab* de deux fonctions `addlig` (d'arguments `b, i, j, A`) et `echlig` (d'arguments `i,j,A`) permettant d'effectuer respectivement les deux autres opérations élémentaires sur les lignes d'une matrice :
$$L_i \leftarrow L_i + b_i (i \neq j) \quad \text{et} \quad L_i \leftrightarrow L_j (i \neq j).$$
   (b) Expliquer pourquoi la fonction `multligmat` suivante retourne le même résultat `B` que la fonction `multlig`.

```
function B = multligmat(a,i,A)
        [n,p] = size(A)
        D = eye(n,n)
        D(i,i) = a
        B = D*A
endfunction
```

3. Dans cette question, on note n un entier supérieur ou égal à 2 et M une matrice de $\mathcal{M}_n(\mathbf{R})$ de rang 1.
Pour tout couple $(i, j) \in [\![1, n]\!]^2$, on note $E_{i,j}$ la matrice de $\mathcal{M}_n(\mathbf{R})$ dont tous les coefficients sont nuls sauf celui situé à l'intersection de sa i-ème ligne et de sa j-ème colonne, qui vaut 1.
   (a) i. Justifier l'existence d'une matrice-colonne non nulle $C = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathcal{M}_{n,1}(\mathbf{R})$ et d'une matrice-ligne non nulle $L = [\ell_1, \ldots, \ell_n] \in \mathcal{M}_{1,n}(\mathbf{R})$ telles que $M = CL$.
   ii. Calculer la matrice MC et en déduire une valeur propre de M.
   iii. Montrer que si le réel $\displaystyle\sum_{i=1}^{n} c_i \ell_i$ est différent de 0, alors la matrice M est diagonalisable.
   (b) i. À l'aide de l'égalité $M = CL$, établir l'existence de deux matrices inversibles P et Q telles que $PMQ = E_{1,1}$.
   ii. En déduire que pour tout couple $(i, j) \in [\![1, n]\!]^2$, il existe deux matrices inversibles P_i et Q_j telles que $P_i M Q_j = E_{i,j}$.

Problème

Dans ce problème, on définit et on étudie les fonctions génératrices des moments et les fonctions génératrices des cumulants de variables aléatoires discrètes ou à densité.
Les cumulants d'ordre 3 et 4 permettent de définir des paramètres d'asymétrie et d'aplatissement qui viennent compléter la description usuelle d'une loi de probabilité par son espérance (paramètre de position) et sa variance (paramètre de dispersion) ; ces cumulants sont notamment utilisés pour l'évaluation des risques financiers.
Dans tout le problème :

— on note $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé et toutes les variables aléatoires introduites dans l'énoncé sont des variables aléatoires réelles définies sur $(\Omega, \mathcal{A})$;

— sous réserve d'existence, l'espérance et la variance d'une variable aléatoire X sont respectivement notées $\mathbf{E}[X]$ et $\mathbf{V}(X)$;

— pour toute variable aléatoire X et pour tout réel t pour lesquels la variable aléatoire e^{tX} admet une espérance, on pose :

$$M_X(t) = \mathbf{E}\left[e^{tX}\right] \quad \text{et} \quad K_X(t) = \ln(M_X(t)) \ ;$$

(les fonctions M_X et K_X sont respectivement appelées la *fonction génératrice des moments* et la *fonction génératrice des cumulants* de X)

— lorsque, pour un entier $p \in \mathbf{N}^*$, la fonction K_X est de classe $\mathcal{C}^p$ sur un intervalle ouvert contenant l'origine, on appelle *cumulant d'ordre* p de X, noté $Q_p(X)$, la valeur de la dérivée p-ème de K_X en 0 :

$$Q_p(X) = K_X^{(p)}(0).$$

Partie I. Fonction génératrice des moments de variables aléatoires discrètes

Dans toute cette partie :

— on note n un entier supérieur ou égal à 2 ;

— toutes les variables aléatoires considérées sont discrètes et à valeurs entières ;

— on note S une variable aléatoire à valeurs dans $\{-1, 1\}$ dont la loi est donnée par :

$$\mathbf{P}(S = -1) = \mathbf{P}(S = 1) = \frac{1}{2}.$$

1. Soit X une variable aléatoire à valeurs dans $[\![-n, n]\!]$.
   (a) Pour tout $t \in \mathbf{R}$, écrire $M_X(t)$ sous la forme d'une somme et en déduire que la fonction M_X est de classe $\mathcal{C}^\infty$ sur $\mathbf{R}$.
   (b) Justifier pour tout $p \in \mathbf{N}^*$, l'égalité : $M_X^{(p)}(0) = \mathbf{E}[X^p]$.
   (c) Soit Y une variable aléatoire à valeurs dans $[\![-n, n]\!]$ dont la fonction génératrice des moments M_Y est la même que celle de X.
   On note G_X et G_Y les deux polynômes définis par :

$$\forall x \in \mathbf{R}, \quad \begin{cases} G_X(x) &= \displaystyle\sum_{k=0}^{2n} \mathbf{P}(X = k - n)\, x^k \\ G_Y(x) &= \displaystyle\sum_{k=0}^{2n} \mathbf{P}(Y = k - n)\, x^k. \end{cases}$$

   i. Vérifier pour tout $t \in \mathbf{R}$, l'égalité $G_X(e^t) = e^{nt} M_X(t)$.
   ii. Justifier la relation : $\forall t \in \mathbf{R}$, $G_X(e^t) = G_Y(e^t)$.
   iii. En déduire que la variable aléatoire Y suit la même loi que X.

2. Dans cette question, on note X_2 une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(2, \frac{1}{2}\right)$.
   On suppose que les variables aléatoire X_2 et S sont indépendantes et on pose $Y_2 = SX_2$.
   (a) i. Préciser l'ensemble des valeurs possibles de la variable aléatoire Y_2.
   ii. Calculer les probabilités $\mathbf{P}(Y_2 = y)$ attachées au diverses valeurs possibles y de Y_2.

(b) Vérifier que la variable aléatoire $X_2 - (S+1)$ suit la même loi que Y_2.

3. Le script *Scilab* suivant permet d'effectuer des simulations de la variable aléatoire Y_2 définie dans la question précédente.

```
(1)        n = 10
(2)       X = grand(n,2,'bin',2,0.5)
(3)       B = grand(n,2,'bin',1,0.5)
(4)       S = 2*B-ones(n,2)
(5)       Z1 = [S(1:n,1).*X(1:n,1),X(1:n,1)-S(1:n,1)-ones(n,1)]
(6)        Z2 = [S(1:n,1).*X(1:n,1),X(1:n,2)-S(1:n,2)-ones(n,1)]
```

(a) Que contiennent les variables `X` et `S` après l'exécution des quatre premières instructions ?

(b) Expliquer pourquoi, après l'exécution des six instructions, chacun des coefficients des matrices `Z1` et `Z2` contient une simulation de la variable aléatoire Y_2.

(c) On modifie la première ligne du script précédent en affectant à `n` une valeur beaucoup plus grande que 10 (par exemple 100000) et en lui adjoignant les deux instructions `(7)` et `(8)` suivantes :

```
(7)       p1 = length(find(Z1(1:n,1)==Z1(1:n,2)))/n
(8)       p2 = length(find(Z2(1:n,1)==Z2(1:n,2)))/n
```

Quelles valeurs numériques approchées la loi faible des grands nombres permet-elle de fournir pour `p1` et `p2` après l'exécution des huit lignes du nouveau script ?

Dans le langage *Scilab*, la fonction `length` fournit la « longueur »d'un vecteur ou d'une matrice et la fonction `find` calcule les positions des coefficients d'une matrice pour lesquels une propriété est vraie, comme l'illustre le script suivant :

```
--> A = [1;2;0;4]
--> B = [2;2;4;3]
--> length(A)
ans = 4.
--> length([A,B])
ans = 8.
--> find(A<B)
ans = 1. 3. //car 1<2 et 0<4, alors que 2 >=2 et 4>=3
```

4. Dans cette question, on note X_n une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(n, \dfrac{1}{2}\right)$.

On suppose que les variables aléatoires X_n et S sont indépendantes et on pose $Y_n = SX_n$.

(a) Justifier que la fonction M_{X_n} est définie sur $\mathbf{R}$ et calculer $M_{X_n}(t)$ pour tout $t \in \mathbf{R}$.

(b) Montrer que la fonction M_{Y_n} est donnée par : $\forall t \in \mathbf{R}$, $M_{Y_n}(t) = \dfrac{1}{2^{n+1}}\left((1+e^t)^n + (1+e^{-t})^n\right)$.

(c) En utilisant l'égalité $(1+e^{-t})^n = e^{-nt}(1+e^t)^n$, montrer que Y_n suit la même loi que la différence $X_n - H_n$, où H_n est une variable aléatoire indépendante de X_n dont on précisera la loi.

Partie II. Propriétés générales des fonctions génératrices des cumulants et quelques exemples

5. Soit X une variable aléatoire et $\mathcal{D}_X$ le domaine de définition de la fonction K_X.

(a) Donner la valeur de $K_X(0)$.

(b) Soit $(a,b) \in \mathbf{R}^2$ et $Y = aX + b$. Justifier pour tout réel t pour lequel at appartient à $\mathcal{D}_X$, l'égalité :

$$K_Y(t) = bt + K_X(at).$$

(c) On suppose ici que les variables aléatoires X et $-X$ suivent la même loi.

Que peut-on dire dans ce cas des cumulants d'ordre impair de la variable aléatoire X ?

6. Soit X et Y deux variables aléatoires indépendantes et $\mathcal{D}_X$ et $\mathcal{D}_Y$ les domaines de définition respectifs des fonctions K_X et K_Y.

(a) Montrer que pour tout réel t appartenant à la fois à $\mathcal{D}_X$ et $\mathcal{D}_Y$, on a : $K_{X+Y}(t) = K_X(t) + K_Y(t)$.

(b) En déduire une relation entre les cumulants des variables aléatoires X, Y et $X+Y$.

7. Soit U une variable aléatoire suivant la loi uniforme sur l'intervalle $[0,1]$.

(a) Montrer que la fonction M_U est définie sur $\mathbf{R}$ et donnée par : $\forall t \in \mathbf{R}$, $M_U(t) = \begin{cases} \dfrac{e^t - 1}{t} & \text{si} \quad t \neq 0 \\ 1 & \text{si} \quad t = 0. \end{cases}$

(b) Calculer la dérivée de la fonction M_U en tout point $t \neq 0$.

(c) Trouver la limite du quotient $\dfrac{M_U(t) - 1}{t}$ lorsque t tend vers 0.

(d) Montrer que la fonction M_U est de classe $\mathcal{C}^1$ sur $\mathbf{R}$.

8. Soient α et β deux réels tels que $\alpha < \beta$.
Dans cette question, on note X une variable aléatoire qui suit la loi uniforme sur l'intervalle $[\alpha, \beta]$.

(a) Exprimer K_X en fonction de M_U, où la variable aléatoire U a été définie dans la question 7.

(b) Justifier que la fonction K_X est de classe $\mathcal{C}^1$ sur $\mathbf{R}$ et établir l'égalité : $Q_1(X) = \mathbf{E}[X]$.

9. Soit un réel $\lambda > 0$ et soit T une variable aléatoire qui suit la loi de Poisson de paramètre λ.

(a) Déterminer les fonctions M_T et K_T.

(b) En déduire tous les cumulants de T.

10. Soit Z une variable aléatoire qui suit la loi normale centrée réduite.

(a) Justifier pour tout $t \in \mathbf{R}$, la convergence de l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \exp\left(tx - \frac{x^2}{2}\right) dx$.

(b) Montrer que la fonction M_Z est définie sur $\mathbf{R}$ et donnée par : $\forall t \in \mathbf{R}$, $M_Z(t) = \exp\left(\dfrac{t^2}{2}\right)$.

(c) En déduire la valeur de tous les cumulants d'une variable aléatoire qui suit une loi normale d'espérance $\mu \in \mathbf{R}$ et d'écart-type $\sigma \in \mathbf{R}_+^*$.

11. Soit $(T_n)_{n \in \mathbf{N}^*}$ une suite de variables aléatoires telles que, pour tout $n \in \mathbf{N}^*$, la variable aléatoire T_n suit la loi de Poisson de paramètre n. Pour tout $n \in \mathbf{N}^*$, on pose $W_n = \dfrac{T_n - n}{\sqrt{n}}$.

(a) Justifier la convergence en loi de la suite de variables aléatoires $(W_n)_{n \in \mathbf{N}^*}$ vers une variable aléatoire W.

(b) Déterminer la fonction K_{W_n}.

(c) Montrer que pour tout $t \in \mathbf{R}$, on a : $\lim\limits_{n \to +\infty} K_{W_n}(t) = K_W(t)$.

Partie III. Cumulant d'ordre 4.

Dans cette partie, on considère une variable aléatoire X telle que M_X est de classe $\mathcal{C}^4$ sur un intervalle ouvert I contenant l'origine.
On *admet* alors que X possède des moments jusqu'à l'ordre 4 qui coïncident avec les dérivées successives de la fonction M_X en 0. Autrement dit, pour tout $k \in [\![1, 4]\!]$, on a $M_X^{(k)}(0) = \mathbf{E}\left[X^k\right]$.
De plus, on pose : $\mu_4(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^4\right]$.

12. Justifier les égalités : $Q_1(X) = \mathbf{E}[X]$ et $Q_2(X) = \mathbf{V}(X)$.

13. Soient X_1 et X_2 deux variables aléatoires indépendantes et de même loi que X. On pose $S = X_1 - X_2$.

(a) Montrer que la variable aléatoire S possède un moment d'ordre 4 et établir l'égalité :

$$\mathbf{E}\left[S^4\right] = 2\mu_4(X) + 6(\mathbf{V}(X))^2.$$

(b) Montrer que les fonctions M_S et K_S sont de classe $\mathcal{C}^4$ sur I et que pour tout $t \in I$, on a :

$$M_S^{(4)}(t) = K_S^{(4)}(t)M_S(t) + 3K_S^{(3)}(t)M_S'(t) + 3K_S''(t)M_S''(t) + K_S'(t)M_S^{(3)}(t).$$

(c) En déduire l'égalité : $\mathbf{E}\left[S^4\right] = Q_4(S) + 3(\mathbf{V}(S))^2$.

14. Justifier que le cumulant d'ordre 4 de X est donné par la relation : $Q_4(X) = \mu_4(X) - 3(\mathbf{V}(X))^2$.