RAPPORT DE CORRECTION

DE MATHÉMATIQUES Option E

Conception HEC Paris – ESSEC BS

SOMMAIRE

Le sujet

BCE BANQUE COMMUNE D'ÉPREUVES

Code sujet : 288

Conception : HEC Paris – ESSEC BS

OPTION ÉCONOMIQUE

MATHÉMATIQUES

Mardi 30 avril 2019, de 14 h. à 18 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.*
*Aucun document n'est autorisé. **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.

EXERCICE

1. Dans cette question, on considère les matrices $C = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbf{R})$, $L = (1 \ \ 2 \ \ -1) \in \mathcal{M}_{1,3}(\mathbf{R})$ et le produit matriciel $M = C\,L$.

   a) (i) Calculer M et M^2.
   (ii) Déterminer le rang de M.
   (iii) La matrice M est-elle diagonalisable ?

   b) (i) Soit $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -2 & 1 \end{pmatrix}$. Justifier que la matrice P est inversible et calculer le produit $P \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

   (ii) Trouver une matrice inversible Q dont la transposée ${}^tQ$ vérifie : ${}^tQ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

   (iii) Pour une telle matrice Q, calculer le produit $P\,M\,Q$.

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 $L_i \leftarrow a\,L_i$ (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 :

$$L_i \leftarrow L_i + b\,L_j \; (i \neq j) \quad \text{et} \quad L_i \leftrightarrow L_j \; (i \neq j)\,.$$

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 $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbf{R})$ et d'une matrice-ligne non nulle $L = (\ell_1 \;\ldots\; \ell_n) \in \mathcal{M}_{1,n}(\mathbf{R})$ telles que $M = C\,L$.

(ii) Calculer la matrice $M\,C$ 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 $M = C\,L$, établir l'existence de deux matrices inversibles P et Q telles que $P\,M\,Q = E_{1,1}$.

(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 $(\Omega, \mathcal{A}, P)$ 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 $E(X)$ et $V(X)$;
- pour toute variable aléatoire X et pour tout réel t pour lesquels la variable aléatoire e^{tX} admet une espérance, on pose :
$$M_X(t) = E(\mathrm{e}^{tX}) \quad \text{et} \quad K_X(t) = \ln\big(M_X(t)\big)\,;$$
(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 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-ième 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 à valeurs entières ;
- on note S une variable aléatoire à valeurs dans $\{-1, 1\}$ dont la loi est donnée par :

$$P\big([S=-1]\big) = P\big([S=+1]\big) = \frac{1}{2}\,.$$

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 C^∞ sur $\mathbf{R}$.

   b) Justifier pour tout $p \in \mathbf{N}^*$, l'égalité : $M_X^{(p)}(0) = E(X^p)$.

   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 :

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

   (*i*) Vérifier pour tout $t \in \mathbf{R}$, l'égalité : $G_X(\mathrm{e}^t) = \mathrm{e}^{nt}M_X(t)$.

   (*ii*) Justifier la relation : $\forall t \in \mathbf{R},\ G_X(\mathrm{e}^t) = G_Y(\mathrm{e}^t)$.

   (*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 aléatoires X_2 et S sont indépendantes et on pose $Y_2 = S\,X_2$.

   a) (*i*) Préciser l'ensemble des valeurs possibles de la variable aléatoire Y_2.

   (*ii*) Calculer les probabilités $P\big([Y_2 = y]\big)$ attachées aux 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 $\mathcal{B}\left(n, \frac{1}{2}\right)$.
   On suppose que les variables aléatoires X_n et S sont indépendantes et on pose $Y_n = S\,X_n$.
   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 : $\forall t \in \mathbf{R},\ M_{Y_n}(t) = \dfrac{1}{2^{n+1}}\left((1+\mathrm{e}^t)^n + (1+\mathrm{e}^{-t})^n\right)$.
   c) En utilisant l'égalité $(1+\mathrm{e}^{-t})^n = \mathrm{e}^{-nt}(1+\mathrm{e}^t)^n$, 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 $a\,t$ appartient à $\mathcal{D}_X$, l'égalité :
   $$K_Y(t) = b\,t + K_X(a\,t)\,.$$
   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 : $\forall t \in \mathbf{R},\ M_U(t) = \begin{cases} \dfrac{\mathrm{e}^t - 1}{t} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0 \end{cases}$.
   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 C^1 sur $\mathbf{R}$.

8. Soit 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 C^1 sur $\mathbf{R}$ et établir l'égalité : $Q_1(X) = E(X)$.

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 $\displaystyle\int_{-\infty}^{+\infty} \exp\left(t\,x - \frac{x^2}{2}\right) \mathrm{d}x$.

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 : $\displaystyle\lim_{n\to+\infty} K_{W_n}(t) = K_W(t)$.

Partie III. Cumulant d'ordre 4

Dans cette partie, on considère une variable aléatoire X telle que M_X est de classe 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 $M_X^{(k)}(0) = E(X^k)$.

De plus, on pose : $\mu_4(X) = E\Big(\big(X - E(X)\big)^4\Big)$.

12. Justifier les égalités : $Q_1(X) = E(X)$ et $Q_2(X) = V(X)$.

13. Soit 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 :

$$E(S^4) = 2\,\mu_4(X) + 6\big(V(X)\big)^2.$$

b) Montrer que les fonctions M_S et K_S sont de classe 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 : $E(S^4) = Q_4(S) + 3\big(V(S)\big)^2$.

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

Le sujet

Le sujet de cette année était composé comme à l'accoutumée d'un exercice et d'un problème.

L'exercice portait sur l'algèbre linéaire (rang, inversibilité, caractère diagonalisable, etc.) et sur des questions de *Scilab* concernant les opérations élémentaires sur les lignes d'une matrice.

Le problème comportait trois parties, et avait pour objet l'étude des fonctions génératrices des moments et des 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 (skweness) et d'aplatissement (kurtosis) qui viennent compléter la description habituelle d'une loi de probabilité par son espérance ou sa médiane (paramètres de position) et sa variance (paramètre de dispersion).

Ces cumulants sont utilisés, notamment dans le domaine de la finance de marché, pour évaluer des risques financiers ; par exemple, dans une distribution leptokurtique, les échantillons ayant des extrémités plus épaisses que la normale impliquent des valeurs anormales plus fréquentes.

Les outils utilisés étaient ceux du calcul des probabilités et de l'analyse.

On trouvait également dans ce problème, des questions de *Scilab* permettant de simuler une certaine variable aléatoire.

Le barème

L'exercice et le problème comptaient respectivement pour 26% et 74% des points de barème.
Plus précisément, les parties I et II du problème représentait chacune, 33% des points de barème, tandis que la pondération de la partie III était de 8 %.

Le poids des questions de *Scilab* était élevé puisqu'il représentait 19% des points de barème et les questions les plus cotées étaient :

- dans l'exercice, les questions 1.a), 1.b) et 3.a) (16 % des points de barème);
- dans le problème, les questions 1, 2, 3, 4, 5, 7 et 10 (50 % des points de barème).

Remarques de correction

Pour ce qui est de la forme, on observe encore trop de copies sales, parfois à la limite de la lisibilité, une écriture anarchique et un non-respect des lignes horizontales. Par contre, une copie claire, propre avec des résultats encadrés se verra attribuer un « bonus » de points pouvant aller jusqu'à 5 points de barème.
Le jury insiste pour que les candidats numérotent les questions, par exemple, par (2.a, 2.b, 2.c…) et non pas (2.a, b, c...)

Pour ce qui concerne le fond, le cours n'est pas maîtrisé : tout semble flou, même au niveau du vocabulaire. Une démonstration se traduit souvent par une vérification. Les questions relatives à l'informatique, la convergence d'intégrale, la notion de limite ou de fonction de classe C^1 sont souvent, quand elles sont abordées, traitées très approximativement et sans rigueur.

Exercice

Les notions de rang et de diagonalisation d'une matrice sont souvent mal dominées. Ainsi, aucun argument n'est fourni pour affirmer qu'une matrice n'ayant qu'une seule valeur propre n'est pas diagonalisable !!
De même, il ne suffit pas de dire que « l'on voit que les trois vecteurs ne sont pas proportionnels » pour conclure que « la famille est libre ».

Problème

Le problème reposait en grande partie sur l'utilisation du théorème de transfert.
Le fait que certaines variables aléatoires prenaient des valeurs négatives a pu perturber certains candidats et a nécessité un travail rigoureux sur les indices ou les quantificateurs, ce qui a permis de distinguer les candidats ayant un recul suffisant sur ces notions et faisant preuve d'une conduite rigoureuse des calculs, de ceux qui répétaient machinalement les recettes apprises en classe.
On observe un certain nombre d'erreurs grossières telles que :

- $\exp(ab) = \exp(a)\exp(b)$;
- $\ln(E(\exp(tX))) = E(\ln(\exp(tX)))$;
- $E(\exp(tX)) = \exp(E(tX))$;
- $E(XY) = E(X)\,E(Y)$ par linéarité ;
- U équivalent à V implique $\exp(U)$ équivalent à $\exp(V)$.

On rappelle également que $\ln(ab) = \ln(a) + \ln(b)$ si et seulement si $a > 0$ et $b > 0$.
La majorité des candidats ne savent pas étudier la convergence d'une intégrale et il y a encore trop de candidats qui écrivent que l'intégrale d'un produit est égale au produit des intégrales !!!
Très peu de candidats savent utiliser un développement limité.
Enfin, le théorème limite central est inconnu chez nombre de candidats.

Conseils aux futurs candidats

Pour ce qui concerne la forme, le jury conseille aux futurs candidats de lire attentivement le texte préliminaire qui précède toute épreuve écrite de mathématiques, dans lequel il est précisé notamment, que la lisibilité et la qualité de la rédaction entrent pour une part non négligeable dans l'appréciation des copies : un correcteur ne s'attarde pas à essayer de « décrypter » une copie illisible. Par contre, une copie propre et claire ne peut qu'avantager son auteur. Le jury rappelle également que les abréviations dans les copies doivent être proscrites et il conseille de bien numéroter les questions et d'encadrer les résultats.

De plus, les raisonnements doivent être clairs et précis, les affirmations étant étayées par une argumentation solide. Par exemple, le recours trop fréquent à des phrases du type « il est clair que… » doit être évité au profit d'une justification correcte fondée sur un apprentissage rigoureux et une très bonne maîtrise du cours.

Le jury recommande aux futurs candidats de prendre le temps de lire l'ensemble du sujet, non seulement pour s'en imprégner, mais aussi pour pointer les questions qui paraissent faciles à résoudre, lesquelles ne se situent pas nécessairement dans la première partie du sujet.

La recherche d'une solution à une question ne doit pas dépasser quatre à cinq minutes. Au-delà de ce délai, en cas d'échec, le candidat doit admettre le résultat de cette question (si la réponse figure dans l'énoncé), passer à la question suivante sans éprouver un sentiment de déstabilisation ou de découragement. Autrement dit, le jury recommande aux futurs candidats de faire preuve d'une grande ténacité.

Statistiques

Sur les 2212 candidats ayant composé dans cette épreuve, la note moyenne est de 9,56 avec un écart-type de 5,12, ces statistiques étant très voisines de celles du concours 2018.
Environ, 9%, soit 193 candidats, obtiennent une note supérieure à 16 et 23 candidats se voient attribuer la note maximale de 20. La note médiane est de 10,6 et les premier et troisième quartiles sont égaux à 5,0 et 13,6 respectivement.
La note de 20 a été attribuée aux 23 candidats qui ont obtenu au moins un peu plus de la moitié des points du barème.