Épreuve Composite

E C G

DS n°3

Option économique

MATHEMATIQUES

30 Novembre 2024

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 une erreur d'énoncé, il le signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.
Les questions précédées de (*) sont réservées aux cubes.

Exercice

Soit n et p deux entiers supérieurs ou égaux à 1. Si M est une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$, la matrice ${}^tM$ de $\mathcal{M}_{p,n}(\mathbb{R})$ désigne la transposée de M.
On identifie les ensembles $\mathcal{M}_{1,1}(\mathbb{R})$ et $\mathbb{R}$ en assimilant une matrice de $\mathcal{M}_{1,1}(\mathbb{R})$ à son unique coefficient.
On note $\mathcal{B}_n$ la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ et $\mathcal{B}_p$ la base canonique de $\mathcal{M}_{p,1}(\mathbb{R})$.
Si $M \in \mathcal{M}_{n,p}(\mathbb{R})$ et $N \in \mathcal{M}_{p,q}(\mathbb{R})$ ($q \in \mathbb{N}^*$), *on admet* que ${}^t(MN) =^t N\,{}^tM$.

1. Soit X une matrice colonne non nulle donnée de $\mathcal{M}_{n,1}(\mathbb{R})$ de composantes $x_1, x_2, ..., x_n$ dans la base $\mathcal{B}_n$.
   On pose : $A = X\,{}^tX$ et $\alpha = {}^tXX$.
   (a) Exprimer A et α en fonction de $x_1, x_2, ..., x_n$.
   (b) (*) Justifier que la matrice A est diagonalisable.
   (c) Soit f l'endomorphisme de $\mathcal{M}_{n,1}(\mathbb{R})$ de matrice A dans la base $\mathcal{B}_n$. Déterminer $\mathrm{Im}f$ et $\mathrm{Ker}f$, donner une base de Imf et préciser la dimension de $Kerf$.
   (d) Calculer la matrice AX.
   (e) (*) Déterminer les valeurs propres de A ainsi que les sous-espaces propres associés.
2. On suppose que n et p vérifient $1 \leq p \leq n$. Soit $(V_1, V_2, ..., V_p)$ une famille libre de p vecteurs de $\mathcal{M}_{n,1}(\mathbb{R})$.
   On note V la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ dont les colonnes sont, dans cet ordre, $V_1, V_2, ..., V_p$.
   Soit g l'application linéaire de $\mathcal{M}_{p,1}(\mathbb{R})$ dans $\mathcal{M}_{n,1}(\mathbb{R})$ de matrice V dans les bases $\mathcal{B}_p$ et $\mathcal{B}_n$.
   (a) Justifier que le rang de V est égal à p. Déterminer $\mathrm{Ker}g$.
   (b) Soit Y une matrice colonne de $\mathcal{M}_{p,1}(\mathbb{R})$.
   Montrer que l'on a $VY = 0$ si et seulement si l'on a ${}^tVVY = 0$.
   (c) En déduire que la matrice ${}^tVV$ est inversible.

Problème

Dans tout le problème :

- toutes les variables aléatoires introduites sont supposées définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$;
- on note n un entier supérieur ou égal à 2 .

L'objet du problème est l'étude de sommes de variables aléatoires suivant une loi de Bernoulli de même paramètre, mais qui ne sont pas nécessairement indépendantes.
Les parties II et III sont indépendantes de la partie I.

Partie I. Valeurs possibles du coefficient de corrélation linéaire dans divers schémas de Bernoulli

Dans cette partie, on considère des variables aléatoires $X_1, X_2, \ldots, X_n$ suivant chacune la même loi de Bernoulli de paramètre p avec $0 < p < 1$, c'est à dire : $\forall k \in [\![1, n]\!], P([X_k = 1]) = p$ et $P([X_k = 0]) = 1 - p$.
On suppose que pour tout couple $(k, \ell) \in [\![1, n]\!]^2$ avec $k \neq \ell$, le coefficient de corrélation linéaire des variables X_k et X_ℓ est le même ; on note r ce coefficient. On a donc :

$$\forall (k, \ell) \in [\![1, n]\!]^2, \quad \frac{\operatorname{Cov}(X_k, X_\ell)}{\sqrt{V(X_k) V(X_\ell)}} = \begin{cases} 1 & \text{si } k = \ell \\ r & \text{si } k \neq \ell \end{cases}$$

1. (a) Dans les cas (i) et (ii) suivants, calculer la valeur de r et exprimer la variance de la variable aléatoire $\sum_{k=1}^{n} X_k$ en fonction de n et p.
      i. Les variables aléatoires $X_1, X_2, \ldots, X_n$ sont mutuellement indépendantes.
      ii. Les variables aléatoires $X_1, X_2, \ldots, X_n$ sont toutes égales.

   De plus, préciser la loi de $\sum_{k=1}^{n} X_k$ dans chacun des deux cas précédents.

   (b) Montrer que pour tout $k \in [\![1, n]\!]$, la variance de la variable aléatoire $\sum_{i=1}^{k} X_i$ est donnée par la formule :

   $$V\left(\sum_{i=1}^{k} X_i\right) = kp(1-p)(1 + (k-1)r)$$

   (c) En déduire que le coefficient r est au moins égal à $-\frac{1}{n-1}$.

2. On suppose dans cette question que n est au moins égal à 2 .
   (a) Montrer que r est égal à -1 si et seulement si on a : $P([X_1 = 1] \cap [X_2 = 1]) = p(2p - 1)$.

   (b) Que vaut alors $P([X_1 = 0] \cap [X_2 = 0])$?

   (c) En déduire que le coefficient r ne peut-être égal à -1 que lorsque $p = \frac{1}{2}$ et $P([X_1 + X_2 = 1]) = 1$.

3. On suppose dans cette question que n est supérieur ou égal à 3 et que $P\left(\left[\sum_{k=1}^{n} X_k = 1\right]\right) = 1$.

   (a) Exprimer les valeurs de p et r en fonction de n.

   (b) Déterminer les n-uplets $(x_1, x_2, \ldots, x_n) \in \{0, 1\}^n$ pour lesquels la probabilité $P\left(\bigcap_{k=1}^{n} [X_k = x_k]\right)$ est strictement positive et la calculer.

Partie II. Lois bêtas-binomiales

4. Soit $(x, y) \in \mathbf{R}^2$.

   (a) Justifier que l'intégrale $\displaystyle\int_0^{\frac{1}{2}} t^{x-1}(1-t)^{y-1}dt$ est convergente si seulement si $x > 0$.

   (b) Pour tout réel ε tel que $0 < \varepsilon < \dfrac{1}{2}$, établir à l'aide d'un changement de variable affine, l'égalité

$$\int_{\frac{1}{2}}^{1-\varepsilon} t^{x-1}(1-t)^{y-1}dt = \int_{\varepsilon}^{\frac{1}{2}} t^{y-1}(1-t)^{x-1}dt$$

   (c) En déduire que l'intégrale $\displaystyle\int_0^1 t^{x-1}(1-t)^{y-1}dt$ est convergente si et seulement si $x > 0$ et $y > 0$.

Dans toute la suite du problème, on pose : $\quad \forall (x, y) \in \left(\boldsymbol{R}^{+*}\right)^2, \quad B(x, y) = \displaystyle\int_0^1 t^{x-1}(1-t)^{y-1}dt.$

5. Soit x et y des réels strictement positifs.

   (a) À l'aide d'une intégration par parties, établir la relation : $\quad B(x+1, y) = \dfrac{x}{y} \times B(x, y+1).$

   (b) En déduire l'égalité : $\quad B(x, y+1) = \dfrac{y}{x+y} \times B(x, y).$

6. Pour tout réel z, soit $\left((z)^{[m]}\right)_{m \in \mathbf{N}}$ la suite définie par : $(z)^{[0]} = 1$ et $\forall m \in \mathbf{N}, (z)^{[m+1]} = (z+m) \times (z)^{[m]}$.
   (par exemple, pour tout $m \in \mathbf{N}$, on a $(1)\ ^{[m]} = m!$)
   Établir pour tout $(x, y) \in \left(\mathbf{R}^{+*}\right)^2$ et pour tout couple (k, ℓ) d'entiers tels que $0 \leqslant k \leqslant \ell$, la relation :

$$B(x+k, y+\ell-k) = \frac{(x)^{[k]} \times (y)^{[\ell-k]}}{(x+y)^{[\ell]}} \times B(x, y)$$

7. Soit a et b des réels strictement positifs.
   Pour $k \in [[0, n]$, on pose : $\quad p_k = \dbinom{n}{k} \dfrac{(a)^{[k]} \times (b)^{[n-k]}}{(a+b)^{[n]}}.$

   (a) À l'aide de la relation obtenue dans la question 6, montrer que $\displaystyle\sum_{k=0}^{n} p_k = 1.$

   On dit qu'une variable aléatoire S suit une loi bêta-binomiale $\boldsymbol{B}(n; a, b)$ si $S(\Omega) = [\![0, n]\!]$ et si :

$$\forall k \in \Big[[0, n], \quad P(S = k) = \binom{n}{k} \frac{(a)^{[k]} \times (b)^{[n-k]}}{(a+b)^{[n]}}$$

   (b) Reconnaître la loi $\boldsymbol{B}(n; 1, 1)$.

   (c) Montrer que l'espérance d'une variable aléatoire S qui suit la loi $\boldsymbol{B}(n; a, b)$ est égale à $\dfrac{na}{a+b}$.

Partie III. Un modèle possible dans le cas où $n = 2$

Soit a et b des réels strictement positifs et X_1 et X_2 deux variables aléatoires à valeurs dans $\{0, 1\}$ telles que :

$$\forall (x_1, x_2) \in \{0, 1\}^2 \quad P\left([X_1 = x_1] \cap [X_2 = x_2]\right) = \frac{B\left(a + x_1 + x_2, b + 2 - x_1 - x_2\right)}{B(a, b)}$$

8. (a) Montrer que les deux variables X_1 et X_2 suivent la même loi de Bernoulli.

   (b) Montrer que la variable aléatoire $X_1 + X_2$ suit la loi bêta-binomiale $\boldsymbol{B}(2; a, b)$.

(c) Établir la relation : $P_{[X_1=1]}\left([X_2=1]\right) = \dfrac{a+1}{a+b+1}$.

9. La fonction Python suivante dont le script est incomplet (lignes (5) et (6)), effectue une simulation des deux variables X_1 et X_2 qu'elle place dans un vecteur ligne à deux composantes.

```
1  def randbetabin(a,b):
2          x=np.zeros([1,2])
3          u=(a+b)*np.rand()
4          v=(a+b+1)*np.rand()
5          if u<a :
6                  x[0][0]=1
7                  if ... :
8                          x[0][1]=1
9          elif ... :
10                         x[0][1]=1
11         return(x)
```

(a) Préciser la loi simulée par la variable u de la ligne (3).

(b) Compléter les lignes (5) et (6).

10. (a) Calculer le coefficient de corrélation linéaire de X_1 et X_2.

(b) Soit (p,r) un couple de réels vérifiant $0<p<1$ et $0<r<1$.
Expliquer comment utiliser la fonction `randbetabin` pour simuler deux variables aléatoires suivant une même loi de Bernoulli de paramètre p et dont le coefficient de corrélation linéaire est égal à r.