



PRÉPASIMPLE

prepasimple.fr

HEC — 2014

Sujet de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Sujet

Dans tout le problème, k désigne un entier supérieur ou égal à 2.

Notations algébriques

Pour tout $n \in \mathbb{N}^*$, on note $\mathcal{M}_{n,1}(\mathbb{R})$ l'ensemble des matrices-colonnes à n lignes à coefficients réels et $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées à n lignes et n colonnes à coefficients réels. On identifie les ensembles $\mathcal{M}_1(\mathbb{R})$ et $\mathbb{R}$ en assimilant une matrice de $\mathcal{M}_1(\mathbb{R})$ à son unique coefficient.

La base canonique de $\mathcal{M}_{k,1}(\mathbb{R})$ est notée $\mathcal{C}_k = (e_1, e_2, \dots, e_k)$ et l'espace vectoriel $\mathcal{M}_{k,1}(\mathbb{R})$ est muni de sa structure euclidienne usuelle pour laquelle la base $\mathcal{C}_k$ est orthonormale. On note $\langle u, v \rangle$ le produit scalaire de deux vecteurs u et v de $\mathcal{M}_{k,1}(\mathbb{R})$ et $\|u\| = \sqrt{\langle u, u \rangle}$ la norme du vecteur u .

Pour toute matrice-colonne d de $\mathcal{M}_{n,1}(\mathbb{R})$ de composantes $d_1, d_2, \dots, d_n$, on note $\text{Diag}(d)$ la matrice diagonale de $\mathcal{M}_n(\mathbb{R})$ définie par :

$$\text{Diag}(d) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}.$$

La transposée d'une matrice M est notée tM et I_k désigne la matrice identité de $\mathcal{M}_k(\mathbb{R})$.

Notations probabilistes

Toutes les variables aléatoires et tous les vecteurs aléatoires qui interviennent dans ce problème sont définis sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

On dit qu'un vecteur aléatoire discret $(Y_1, Y_2, \dots, Y_k)$, à valeurs dans $\mathbb{R}^k$, admet une espérance lorsque chacune de ses composantes en admet une.

On note Y la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ de composantes $Y_1, Y_2, \dots, Y_k$ et $\mathcal{E}(Y)$ la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ dont les composantes sont les espérances $E(Y_1), E(Y_2), \dots, E(Y_k)$.

Lorsque chacune des composantes Y_i ($i \in \llbracket 1, k \rrbracket$) admet une variance, on appelle matrice de variance-covariance de Y , notée $\mathcal{V}(Y)$, la matrice symétrique de $\mathcal{M}_k(\mathbb{R})$ dont les coefficients diagonaux sont les variances $V(Y_i)$ et les coefficients non diagonaux les covariances $\text{Cov}(Y_i, Y_j)$ pour tout $(i, j) \in \llbracket 1, k \rrbracket^2$ avec $i \neq j$.

En résumé, on pose sous réserve d'existence :

$$\mathcal{E}(Y) = \begin{pmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_k) \end{pmatrix}$$

et

$$\mathcal{V}(Y) = \begin{pmatrix} V(Y_1) & \text{Cov}(Y_1, Y_2) & \cdots & \text{Cov}(Y_1, Y_k) \\ \text{Cov}(Y_2, Y_1) & V(Y_2) & \cdots & \text{Cov}(Y_2, Y_k) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(Y_k, Y_1) & \text{Cov}(Y_k, Y_2) & \cdots & V(Y_k) \end{pmatrix}.$$

Dans tout le problème, on note $p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{pmatrix}$ une matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ vérifiant

$$\sum_{i=1}^k p_i = 1$$

et pour tout $i \in \llbracket 1, k \rrbracket$, $p_i \geq 0$.

L'objet du problème est l'étude des propriétés des matrices de variance-covariance en liaison avec la loi des vecteurs aléatoires correspondants.

Partie I. Lois généralisées de Bernoulli

Dans cette partie, on note u la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ dont tous les coefficients valent 1.

1. Soit $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix}$ une matrice-colonne non nulle de $\mathcal{M}_{k,1}(\mathbb{R})$ et

$$\alpha = \sum_{i=1}^k a_i.$$

On pose : $M = a {}^t u$.

- a) Calculer la matrice M et préciser son rang.
- b) Calculer la matrice Ma et en déduire une valeur propre de M .
- c) Montrer que $M^2 = \alpha M$. Que peut-on en déduire sur les valeurs propres de M ?
- d) Montrer que M est diagonalisable si et seulement si $\alpha \neq 0$.
- e) Pour quelles valeurs de α la matrice $I_k - M$ est-elle inversible?
- f) On suppose que $\alpha = 1$. Montrer que M est la matrice dans la base canonique de $\mathbb{R}^k$ d'un projecteur dont on précisera l'image et le noyau. Dans quel cas ce projecteur est-il orthogonal?

On dit qu'un vecteur aléatoire $(X_1, X_2, \dots, X_k)$ suit la loi généralisée de Bernoulli de paramètre p , notée $\mathcal{B}_k(p)$, si on a :

$$\forall i \in \llbracket 1, k \rrbracket, P([X = e_i]) = p_i, \text{ avec } X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix}.$$

2. Soit $(X_1, X_2, \dots, X_k)$ un vecteur aléatoire suivant la loi $\mathcal{B}_k(p)$.

a) Pour $i \in \llbracket 1, k \rrbracket$, comparer les événements $[X = e_i]$ et $[X_i = 1]$; en déduire que chaque variable aléatoire X_i suit une loi de Bernoulli de paramètre p_i et écrire la matrice $\mathcal{E}(X)$.

b) Quelle est la loi de la variable aléatoire $X_1 + X_2$?

c) Montrer que $\text{Cov}(X_1, X_2) = -p_1 p_2$.

d) Écrire la matrice $\mathcal{V}(X)$.

3. Soit $M(p)$ la matrice de $\mathcal{M}_k(\mathbb{R})$ définie par : $M(p) = p {}^t u$.

a) Vérifier l'égalité : $\mathcal{V}(X) = (I_k - M(p)) \text{Diag}(p)$.

b) Montrer que si $p_1, p_2, \dots, p_k$ sont différents de 0, le rang de $\mathcal{V}(X)$ est égal à $k - 1$.

c) Soit σ une permutation de $\llbracket 1, k \rrbracket$ et p_σ la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ de composantes $p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(k)}$. Montrer que $\mathcal{V}(X)$ est semblable à $(I_k - p_\sigma {}^t u) \text{Diag}(p_\sigma)$.

d) Exprimer le rang de $\mathcal{V}(X)$ en fonction du nombre d'éléments i de $\llbracket 1, k \rrbracket$ pour lesquels on a $p_i \neq 0$.

Partie II. Tirages avec remise dans une population stratifiée

Dans cette partie, on suppose que pour tout $i \in \llbracket 1, k \rrbracket$, on a $p_i > 0$ et que $p_1, p_2, \dots, p_k$ sont les proportions d'individus appartenant aux diverses catégories d'une population statistique scindée en k catégories distinctes.

Pour modéliser une suite illimitée de tirages équiprobables avec remise effectués dans cette population, on utilise des variables aléatoires $X_i^{(n)}$ définies par :

$\forall n \in \mathbb{N}^*, \forall i \in \llbracket 1, k \rrbracket, X_i^{(n)} = 1$ si l'individu extrait au n -ième tirage appartient à la i -ème catégorie, et $X_i^{(n)} = 0$ sinon.

On suppose que les vecteurs aléatoires $(X_1^{(n)}, X_2^{(n)}, \dots, X_k^{(n)})$ ($n \in \mathbb{N}^*$) suivent chacun la loi $\mathcal{B}_k(p)$ (partie I) et sont mutuellement indépendants. Cette indépendance mutuelle signifie que pour tout entier $n \geq 2$ et pour toutes fonctions $\varphi_1, \varphi_2, \dots, \varphi_n$ définies sur $\mathbb{R}^k$ à valeurs réelles, les variables aléatoires $\varphi_1(X_1^{(1)}, X_2^{(1)}, \dots, X_k^{(1)}), \varphi_2(X_1^{(2)}, X_2^{(2)}, \dots, X_k^{(2)}), \dots, \varphi_n(X_1^{(n)}, X_2^{(n)}, \dots, X_k^{(n)})$ sont indépendantes.

Pour tout $n \in \mathbb{N}^*$, on note $X^{(n)}$ la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ de composantes $X_1^{(n)}, X_2^{(n)}, \dots, X_k^{(n)}$ et $S^{(n)}$ la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ de composantes $S_1^{(n)}, S_2^{(n)}, \dots, S_k^{(n)}$, où pour tout $i \in \llbracket 1, k \rrbracket$, on a

$$S_i^{(n)} = \sum_{j=1}^n X_i^{(j)}.$$

4. a) Préciser l'ensemble N_n des matrices-colonnes s de $\mathcal{M}_{k,1}(\mathbb{R})$ pour lesquelles on a $P([S^{(n)} = s]) > 0$.

b) Déterminer les lois respectives des deux variables aléatoires $S_1^{(n)}$ et $S_1^{(n)} + S_2^{(n)}$.
Sont-elles indépendantes?

c) Montrer que $\mathcal{V}(S^{(n)}) = n \mathcal{V}(X^{(1)})$.

5. Soit H un élément de $\mathcal{A}$ vérifiant $0 < P(H) < 1$, $\overline{H}$ l'événement contraire de H et W une variable aléatoire discrète admettant une variance.

- a) Justifier l'existence de $E(W^2|H)$, espérance de W^2 pour la probabilité conditionnelle P_H .
- b) On pose : $V(W|H) = E(W^2|H) - (E(W|H))^2$ (variance de W pour la probabilité conditionnelle P_H).

En utilisant le système complet d'événements $(H, \overline{H})$ et la formule de l'espérance totale pour W et W^2 , établir l'inégalité : $V(W) \geq P(H) V(W|H)$.

6. Pour tout $i \in \llbracket 1, k \rrbracket$, on note T_i le temps d'attente du premier tirage d'un individu de la i -ème catégorie et on note T la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ de composantes $T_1, T_2, \dots, T_k$.

- a) Soit $i \in \llbracket 1, k \rrbracket$. Justifier que la probabilité que T_i soit infini est nulle. Quelle est la loi de T_i ?
- b) On pose : $H_k = \bigcap_{i=1}^{k-1} [T_i = i]$. Calculer $P(H_k)$. Préciser la loi conditionnelle de $T_k - (k-1)$ sachant H_k .

En déduire $E(T_k|H_k)$ et $V(T_k|H_k)$.

- c) En exploitant le résultat de la question 5.b), établir pour tout vecteur $v = (v_1, v_2, \dots, v_k)$ de $\mathbb{R}^k$, l'inégalité :

$$V\left(\sum_{i=1}^k v_i T_i\right) \geq \frac{v_k^2(1-p_k)}{p_k^2} \times \prod_{i=1}^{k-1} p_i.$$

- d) Montrer plus généralement que pour tout $j \in \llbracket 1, k \rrbracket$, on a :

$$V\left(\sum_{i=1}^k v_i T_i\right) \geq \frac{v_j^2(1-p_j)}{p_j^2} \times \prod_{i \in \llbracket 1, k \rrbracket, i \neq j} p_i.$$

Partie III. Support et rang stochastiques d'un vecteur aléatoire

Dans toute cette partie, $(Y_1, Y_2, \dots, Y_k)$ désigne un vecteur aléatoire discret, à valeurs dans $\mathbb{R}^k$, dont chaque composante admet une espérance et une variance. On rappelle que Y est la matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ de composantes $Y_1, Y_2, \dots, Y_k$.

7. On appelle support vectoriel de Y , tout sous-espace vectoriel F de $\mathcal{M}_{k,1}(\mathbb{R})$ tel que $P([Y - \mathcal{E}(Y) \in F]) = 1$. On note $\mathcal{S}(Y)$ l'ensemble des supports vectoriels de Y .
 - a) Justifier l'existence d'un plus petit élément de l'ensemble des éléments de $\mathcal{S}(Y)$. Ce plus petit élément est appelé le rang stochastique de Y et noté $R_s(Y)$.
 - b) Dans quels cas le rang stochastique $R_s(Y)$ est-il nul ?
 - c) Montrer que l'intersection de deux supports vectoriels F_1 et F_2 de Y est un support vectoriel de Y .
 - d) En déduire l'existence d'un unique élément F de $\mathcal{S}(Y)$ tel que la dimension de F soit égale à $R_s(Y)$. L'espace vectoriel F est appelé le support stochastique de Y .
8. Soit u une matrice-colonne de $\mathcal{M}_{k,1}(\mathbb{R})$ de composantes $u_1, u_2, \dots, u_k$.
 - a) Montrer que la variable aléatoire

$$\sum_{i=1}^k u_i Y_i$$

admet une variance, égale à ${}^t u \mathcal{V}(Y) u$.

- b) Établir l'existence d'un unique vecteur $(\lambda_1, \lambda_2, \dots, \lambda_k)$ de $\mathbb{R}^k$ tel que $\mathcal{V}(Y)$ soit semblable à la matrice $\text{Diag}(\lambda)$ et pour lequel $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq 0$ (on note $\text{Diag}(\lambda)$ la matrice diagonale de $\mathcal{M}_k(\mathbb{R})$ de coefficients diagonaux $\lambda_1, \lambda_2, \dots, \lambda_k$).

c) On pose :

$$\|Y - \mathcal{E}(Y)\|^2 = \sum_{i=1}^k (Y_i - E(Y_i))^2.$$

Montrer que

$$E(\|Y - \mathcal{E}(Y)\|^2) = \sum_{i=1}^k \lambda_i.$$

9. Soit $q \in \llbracket 1, k \rrbracket$, F un sous-espace vectoriel de $\mathcal{M}_{k,1}(\mathbb{R})$ de dimension q et $(f^{(1)}, f^{(2)}, \dots, f^{(q)})$ une base orthonormale de F .

a) Soit $\omega \in \Omega$. Justifier l'existence de $Q_F(\omega) = \inf \{\|Y(\omega) - \mathcal{E}(Y) - x\|^2 ; x \in F\}$ et montrer que :

$$\|Y(\omega) - \mathcal{E}(Y)\|^2 = Q_F(\omega) + \sum_{j=1}^q \langle Y(\omega) - \mathcal{E}(Y), f^{(j)} \rangle^2.$$

b) À l'aide de la question 8, établir l'égalité :

$$E(Q_F) = \sum_{i=1}^k \lambda_i - \sum_{j=1}^q {}^t f^{(j)} \mathcal{V}(Y) f^{(j)}.$$

c) Que devient l'égalité précédente lorsque $F = \mathcal{M}_{k,1}(\mathbb{R})$?

10. a) Montrer que pour toute matrice-colonne f de $\mathcal{M}_{k,1}(\mathbb{R})$ vérifiant $\|f\| = 1$, on a : ${}^t f \mathcal{V}(Y) f \leq \lambda_1$.

b) En déduire la borne inférieure de $E(Q_F)$ lorsque F décrit l'ensemble des droites vectorielles de $\mathcal{M}_{k,1}(\mathbb{R})$.

- c) Dans cette question, on suppose que $(Y_1, Y_2, \dots, Y_k)$ suit la loi $\mathcal{B}_k(p)$, où pour tout $i \in \llbracket 1, k \rrbracket$, on a $p_i = \frac{1}{k}$.

Calculer les valeurs propres de $\mathcal{V}(Y)$ et la borne inférieure de $E(Q_F)$ pour l'ensemble des droites vectorielles F de $\mathcal{M}_{k,1}(\mathbb{R})$, puis préciser pour quelle(s) droite(s) cette borne est atteinte.

11. On suppose que le rang r de $\mathcal{V}(Y)$ est non nul. On note F_0 la somme des sous-espaces propres associés aux valeurs propres non nulles de $\mathcal{V}(Y)$ et F un sous-espace vectoriel de $\mathcal{M}_{k,1}(\mathbb{R})$ tel que $F \subset F_0$ et $F \neq F_0$.

- a) Calculer $E(Q_{F_0})$ et en déduire que F_0 est un support vectoriel de Y .
- b) Justifier l'existence d'un vecteur $f^{(r)}$ de F_0 , orthogonal à F et de norme 1.
- c) Montrer que ${}^t f^{(r)} \mathcal{V}(Y) f^{(r)} > 0$ et en déduire que $E(Q_F) \neq 0$.
- d) Montrer que le rang stochastique $R_s(Y)$ de Y est égal à r .

12. Dans cette question, on reprend les définitions et notations de la question 6.

- a) À l'aide de la question 6.d), montrer que le rang stochastique $R_s(T)$ de T est égal à k .

- b) Montrer que pour tout $i \in \mathbb{N}^*$ tel que $P([T_1 = i] \cap [T_2 > i]) > 0$, on a :

$$E(T_1 T_2 \mid [T_1 = i] \cap [T_2 > i]) = i \left(i + \frac{1}{p_2} \right).$$

- c) Établir la relation : $E(T_1 T_2) = \frac{1}{p_1 p_2} - \frac{1}{p_1 + p_2}$.

- d) On note $\Pi = (\pi_{i,j})_{1 \leq i,j \leq k}$ la matrice de $\mathcal{M}_k(\mathbb{R})$ définie par : $\pi_{i,j} = \frac{1 - p_i}{p_i^2}$ si $i = j$ et

$$\pi_{i,j} = -\frac{1}{p_i + p_j} \text{ si } i \neq j.$$

Montrer que la matrice Π est inversible.