
DS5 (version B)

Exercice

Soit n un entier supérieur ou égal à 2 et $\mathcal{B} = (e_1, e_2, \dots, e_n)$ la base canonique de \mathbb{R}^n .

Soit v un vecteur donné de \mathbb{R}^n de coordonnées v_1, v_2, \dots, v_n dans la base \mathcal{B} et qui vérifie $\sum_{i=1}^n v_i = 1$.

Soit f l'application définie sur \mathbb{R}^n qui à tout vecteur $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, associe le vecteur $f(x)$ défini par : $f(x) = x - \left(\sum_{i=1}^n x_i \right) \cdot v$.

1. *a)* Montrer que f est un endomorphisme de \mathbb{R}^n .
b) Montrer que $f \circ f = f$.
2. Déterminer le spectre de f .
3. *a)* Montrer que le vecteur y appartient à l'image de f , notée $\text{Im}(f)$, si et seulement si $f(y) = y$.
b) Montrer que la dimension de $\text{Im}(f)$ est inférieure ou égale à $n - 1$.
c) Montrer que pour tout $i \in \llbracket 1, n - 1 \rrbracket$, on a : $(e_i - e_{i+1}) \in \text{Im}(f)$.
d) En déduire une base et la dimension de $\text{Im}(f)$. Quel est le rang de f ?
4. *a)* Déterminer une base du noyau de f .
b) Quels sont les sous-espaces propres de f ?
c) L'endomorphisme f est-il diagonalisable ?
5. Écrire la matrice M de l'endomorphisme f dans la base canonique de \mathbb{R}^n et la matrice M' de f dans une base de vecteurs propres.

Problème

Dans tout le problème :

- toutes les variables aléatoires introduites sont supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$;
- 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, \dots, X_n suivant chacune la même loi de Bernoulli de paramètre p avec $0 < p < 1$, c'est à dire :

$$\forall k \in \llbracket 1, n \rrbracket, \mathbb{P}([X_k = 1]) = p \quad \text{et} \quad \mathbb{P}([X_k = 0]) = 1 - p$$

On suppose que pour tout couple $(k, \ell) \in \llbracket 1, n \rrbracket^2$ 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 :

$$\forall (k, \ell) \in \llbracket 1, n \rrbracket^2, \frac{\text{Cov}(X_k, X_\ell)}{\sqrt{\mathbb{V}(X_k)\mathbb{V}(X_\ell)}} = \begin{cases} 1 & \text{si } k = \ell \\ r & \text{si } k \neq \ell \end{cases}$$

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, \dots, X_n sont mutuellement indépendantes.

(ii) Les variables aléatoires X_1, X_2, \dots, 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 $k \in \llbracket 1, n \rrbracket$, la variance de la variable aléatoire $\sum_{i=1}^k X_i$ est donnée par la formule :

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

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 : $\mathbb{P}([X_1 = 1] \cap [X_2 = 1]) = p(2p-1)$.

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

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

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

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

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

Partie II. Loïs bêtas-binomiales

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

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

b) Pour tout réel ε tel que $0 < \varepsilon < \frac{1}{2}$, 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 $\int_0^1 t^{x-1} (1-t)^{y-1} dt$ est convergente si et seulement si $x > 0$ et $y > 0$.

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

5. Soit x et y des réels strictement positifs.

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

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

6. Pour tout réel z , soit $((z)^{[m]})_{m \in \mathbb{N}}$ la suite définie par :

$$(z)^{[0]} = 1 \quad \text{et} \quad \forall m \in \mathbb{N}, (z)^{[m+1]} = (z+m) \times (z)^{[m]}$$

(par exemple, pour tout $m \in \mathbb{N}$, on a $(1)^{[m]} = m!$)

Établir pour tout $(x, y) \in (\mathbb{R}_+^*)^2$ et pour tout couple (k, ℓ) d'entiers tels que $0 \leq k \leq \ell$, 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 $k \in \llbracket 0, n \rrbracket$, on pose : $p_k = \binom{n}{k} \frac{(a)^{[k]} \times (b)^{[n-k]}}{(a+b)^{[n]}}$.

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

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

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

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

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

Partie III. Un 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 :

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

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 $\mathbf{B}(2; a, b)$.

c) Établir la relation : $\mathbb{P}_{[X_1=1]}([X_2 = 1]) = \frac{a + 1}{a + b + 1}$.

9. La fonction **Scilab** suivante dont le script est incomplet (lignes 7 et 11), effectue une simulation des deux variables X_1 et X_2 qu'elle place dans un vecteur ligne à deux composantes.

```
1  function x = randbetabin(a, b)
2      x = zeros(1,2)
3      u = (a + b) * rand()
4      v = (a + b + 1) * rand()
5      if (u < a) then
6          x(1,1) = 1
7          if ..... then
8              x(1,2) = 1
9          end
10     else
11         if ..... then
12             x(1,2) = 1
13         end
14     end
15 endfunction
```

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

b) Compléter les lignes 7 et 11.

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 .