

## DS8 (version A)

### EXERCICE

Pour tout  $n \in \mathbb{N}^*$ , on note  $\mathcal{M}_n(\mathbb{R})$  l'ensemble des matrices carrées à  $n$  lignes et  $n$  colonnes à coefficients réels et  $B_n$  l'ensemble des matrices de  $\mathcal{M}_n(\mathbb{R})$  dont tous les coefficients sont égaux à 0 ou à 1.

1. *Exemple 1.* Soit  $A$  la matrice de  $B_2$  définie par :  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .

a) Calculer la matrice  $A^2$ .

- 1 pt :  $A^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = I_2$

b) Quelles sont les valeurs propres de  $A$  ?

- 1 pt : **D'après la question 1.a),  $A^2 - I_2 = 0$ . On en déduit que le polynôme  $P(X) = X^2 - 1 = (X - 1)(X + 1)$  est un polynôme annulateur de  $A$ .**

- 1 pt :  $\text{Sp}(A) \subset \{-1, 1\}$

- 1 pt : **1 est bien valeur propre de  $A$**

- 1 pt : **-1 est bien valeur propre de  $A$**

c) La matrice  $A$  est-elle diagonalisable ?

- 1 pt : **La matrice  $A$  est carrée d'ordre 2 et admet 2 valeurs propres distinctes donc la matrice  $A$  est diagonalisable**

2. *Exemple 2.* Soit  $B$  la matrice de  $B_3$  définie par :  $B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .

On considère les instructions et la sortie (ans) **Scilab** suivantes :

```
1 B = [0,1,0;1,0,0;0,0,1]
2 P = [1,1,0;1,-1,0;0,0,1]
3 inv(P) * B * P
```

```
ans =
  1.  0.  0.
  0. -1.  0.
  0.  0.  1.
```

a) Déduire les valeurs propres de  $B$  de la séquence **Scilab** précédente.

- 1 pt : **Notons  $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  et  $P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . D'après la séquence Scilab,**

**$P^{-1}BP = D$ , d'où  $B = PDP^{-1}$**

- 1 pt : **Ainsi,  $B$  est semblable à la matrice diagonale  $D$ . Elle est donc diagonalisable et ses valeurs propres sont les coefficients diagonaux de  $D$**

- 1 pt :  $\text{Sp}(B) = \{-1, 1\}$

b) Déterminer une base de chacun des sous-espaces propres de  $B$ .

- **1 pt** : La famille  $\mathcal{F} = \left( \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$  constituée des colonnes de la matrice  $P$  est une base de vecteurs propres de  $B$
- **2 pt** :  $E_{-1}(B) = \text{Vect} \left( \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right)$
- **2 pt** :  $E_1(B) = \text{Vect} \left( \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$

3. a) Combien existe-t-il de matrices appartenant à  $B_n$  ?

- **1 pt** : Une matrice  $M$  de  $B_n$  est une matrice de  $\mathcal{M}_n(\mathbb{R})$  dont chaque coefficient vaut soit 0 soit 1. On a 2 choix pour chaque coefficient et il y a  $n^2$  coefficients
- **1 pt** :  $\text{Card}(B_n) = 2^{n^2}$

b) Combien existe-t-il de matrices de  $B_n$  dont chaque ligne et chaque colonne comporte exactement un coefficient égal à 1 ?

- **1 pt** : Une matrice  $M$  de  $B_n$  dont chaque ligne et chaque colonne comporte exactement un coefficient égal à 1 est entièrement déterminée par :
  - × la position du 1 sur la 1<sup>ère</sup> ligne :  $n$  possibilités.  
(on peut le placer sur n'importe quelle colonne)
  - × la position du 1 sur la 2<sup>ème</sup> ligne :  $n - 1$  possibilités.  
(on peut le placer sur n'importe quelle colonne non déjà choisie)
  - × la position du 1 sur la 3<sup>ème</sup> ligne :  $n - 2$  possibilités.  
(on peut le placer sur n'importe quelle colonne non déjà choisie)
  - × ...
  - × la position du 1 sur la  $n^{\text{ème}}$  ligne : 1 possibilité.  
(on peut le placer sur n'importe quelle colonne non déjà choisie)
- **1 pt** : L'ensemble des matrices de  $B_n$  dont chaque ligne et chaque colonne comporte exactement un coefficient égal à 1 comporte exactement  $n!$  éléments

4. Dans cette question,  $n$  est un entier supérieur ou égal à 2.

Soit  $E$  un espace vectoriel de dimension  $n$  et  $u$  un endomorphisme de  $E$ . On note :

- $\text{id}$  l'endomorphisme identité de  $E$  ;
- $F$  le noyau de l'endomorphisme  $(u + \text{id})$  et  $G$  le noyau de l'endomorphisme  $(u - \text{id})$  ;
- $p$  la dimension de  $F$  et  $q$  la dimension de  $G$ .

On suppose que  $u \circ u = \text{id}$ .

a) Justifier que l'image de  $(u - \text{id})$  est incluse dans  $F$ .

- **1 pt** : Soit  $y \in \text{Im}(u - \text{id})$ . Montrons que  $y \in F = \text{Ker}(u + \text{id})$
- **1 pt** : Par définition, il existe  $x \in E$  tel que  $y = (u - \text{id})(x) = u(x) - x$
- **3 pt** :  $(u + \text{id})(y) = 0_E$  (**2 pt pour les justifications**)

b) En déduire l'inégalité :  $p + q \geq n$ .

- **1 pt** :  $\text{Im}(u - \text{id}) \subset \text{Ker}(u + \text{id})$ . Ainsi :  $\dim(\text{Im}(u - \text{id})) \leq \dim(\text{Ker}(u + \text{id})) = p$
- **1 pt** : D'après le théorème du rang,  $\dim(\text{Im}(u - \text{id})) = n - q$

On suppose désormais que  $1 \leq p < q$ .

Soit  $(f_1, f_2, \dots, f_p)$  une base de  $F$  et  $(g_1, g_2, \dots, g_q)$  une base de  $G$ .

c) Justifier que  $(f_1, f_2, \dots, f_p, g_1, g_2, \dots, g_q)$  est une base de  $E$ .

- **1 pt** :  $\dim(\text{Ker}(u + \text{id})) = p \geq 1$ . **En particulier** :  $\text{Ker}(u + \text{id}) \neq \{0_E\}$ . **Ainsi,  $-1$  est valeur propre de  $u$  et  $F = \text{Ker}(u + \text{id}) = E_{-1}(u)$**

- **1 pt** :  $\dim(\text{Ker}(u - \text{id})) = q > 1$ . **En particulier** :  $\text{Ker}(u - \text{id}) \neq \{0_E\}$ . **Ainsi,  $1$  est valeur propre de  $u$  et  $G = \text{Ker}(u - \text{id}) = E_1(u)$**

- **1 pt** : **On a alors :**

× la famille  $(f_1, f_2, \dots, f_p)$  est une base de  $E_{-1}(u)$ .

**En particulier, c'est donc une famille libre de  $E_{-1}(u)$ .**

× la famille  $(g_1, g_2, \dots, g_q)$  est une base de  $E_1(u)$ .

**En particulier, c'est donc une famille libre de  $E_1(u)$ .**

- **1 pt** : La famille  $\mathcal{F} = (f_1, f_2, \dots, f_p, g_1, g_2, \dots, g_q)$  est la concaténation de deux familles libres de sous-espaces propres associés à des valeurs propres distinctes. **Ainsi,  $\mathcal{F}$  est une famille libre de  $E$**

- **1 pt** : **par liberté,  $\text{Card}(\mathcal{F}) = p + q \leq \dim(E) = n$**

- **1 pt** : **Or, d'après la question précédente :  $p + q \geq n$  donc  $p + q = n$**

- **1 pt** : **la famille  $\mathcal{F}$  est libre et  $\text{Card}(\mathcal{F}) = n = \dim(E)$  donc  $\mathcal{F}$  est une base de  $E$**

d) Calculer  $u(g_1 - f_1)$  et  $u(g_1 + f_1)$ .

- **1 pt** :  $u(g_1 - f_1) = g_1 + f_1$

- **1 pt** :  $u(g_1 + f_1) = g_1 - f_1$

e) Trouver une base de  $E$  dans laquelle la matrice de  $u$  appartient à  $B_n$ .

- **1 pt** : **On considère  $\mathcal{B}'$  la famille :**

$$\mathcal{B}' = (g_1 - f_1, g_1 + f_1, g_2 - f_2, g_2 + f_2, \dots, g_p - f_p, g_p + f_p, g_{p+1}, g_{p+2}, \dots, g_q)$$

- **3 pt** :  **$\mathcal{B}'$  est une famille libre**

- **1 pt** :  **$\text{Card}(\mathcal{B}') = n = \dim(E)$  donc  $\mathcal{B}'$  est une base de  $E$**

- **3 pt** : **la matrice de  $u$  dans la base  $\mathcal{B}'$  ne contient que des 1 et des 0**

## PROBLÈME

Les tables de mortalité sont utilisées en démographie et en actuariat pour prévoir l'espérance de vie des individus d'une population. On s'intéresse dans ce problème à un modèle qui permet d'ajuster la durée de vie à des statistiques portant sur les décès observés au sein d'une génération.

Dans tout le problème, on note :

- $a$  et  $b$  deux réels strictement positifs ;
- $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé sur lequel sont définies toutes les variables aléatoires du problème ;
- $G_{a,b}$  la fonction définie sur  $\mathbb{R}_+$  par :  $G_{a,b}(x) = \exp\left(-ax - \frac{b}{2}x^2\right)$ .

### Partie I. Loi exponentielle linéaire

1. a) Montrer que la fonction  $G_{a,b}$  réalise une bijection de  $\mathbb{R}_+$  sur l'intervalle  $]0, 1]$ .

b) Pour tout réel  $y > 0$ , résoudre l'équation d'inconnue  $x \in \mathbb{R}$  :  $ax + \frac{b}{2}x^2 = y$ .

c) On note  $G_{a,b}^{-1}$  la bijection réciproque de  $G_{a,b}$ .

Quelle est, pour tout  $u \in [0, 1[$ , l'expression de  $G_{a,b}^{-1}(1 - u)$  ?

2. a) Justifier la convergence de l'intégrale  $\int_0^{+\infty} G_{a,b}(x) dx$ .

b) Soit  $f$  la fonction définie sur  $\mathbb{R}$  par :  $f(x) = \sqrt{\frac{b}{2\pi}} \times \exp\left(-\frac{1}{2}b\left(x + \frac{a}{b}\right)^2\right)$ .

Montrer que  $f$  est une densité d'une variable aléatoire suivant une loi normale dont on précisera les paramètres (espérance et variance).

c) Soit  $\Phi$  la fonction de répartition de la loi normale centrée réduite.

Déduire de la question 2.b), l'égalité :

$$\int_0^{+\infty} G_{a,b}(x) dx = \sqrt{\frac{2\pi}{b}} \times \exp\left(\frac{a^2}{2b}\right) \times \Phi\left(-\frac{a}{\sqrt{b}}\right)$$

3. Pour tout  $a > 0$  et pour tout  $b > 0$ , on pose :  $f_{a,b}(x) = \begin{cases} (a + bx) \exp\left(-ax - \frac{b}{2}x^2\right) & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$ .

a) Justifier que la fonction  $f_{a,b}$  est une densité de probabilité.

On dit qu'une variable aléatoire suit la loi exponentielle linéaire de paramètres  $a$  et  $b$ , notée  $\mathcal{E}_\ell(a, b)$ , si elle admet  $f_{a,b}$  pour densité.

b) Soit  $X$  une variable aléatoire suivant la loi  $\mathcal{E}_\ell(a, b)$ . À l'aide d'une intégration par parties, justifier que  $X$  admet une espérance  $\mathbb{E}(X)$  telle que :  $\mathbb{E}(X) = \int_0^{+\infty} G_{a,b}(x) dx$ .

4. Soit  $Y$  une variable aléatoire suivant la loi exponentielle de paramètre 1. On pose :  $X = \frac{-a + \sqrt{a^2 + 2bY}}{b}$ .

a) Justifier que pour tout réel  $x \in \mathbb{R}_+$ , on a :  $\mathbb{P}([X \geq x]) = G_{a,b}(x)$ .

b) En déduire que  $X$  suit la loi  $\mathcal{E}_\ell(a, b)$ .

c) On note  $U$  une variable aléatoire suivant la loi uniforme sur  $[0, 1[$ .

Déterminer la loi de la variable aléatoire  $G_{a,b}^{-1}(1 - U)$ .

5. La fonction **Scilab** suivante génère des simulations de la loi exponentielle linéaire.

```

1  function x = grandlinexp(a,b,n)
2      u = rand(n,1)
3      y = .....
4      x = (-a + sqrt(a ^ 2 + 2 * b * y)) / b
5  endfunction
    
```

- a) Quelle est la signification de la ligne de code 2 ?
- b) Compléter la ligne de code 3 pour que la fonction **grandlinexp** génère les simulations désirées.

6. De quel nombre réel peut-on penser que les six valeurs générées par la boucle **Scilab** suivante fourniront des valeurs approchées de plus en plus précises et pourquoi ?

```

1  for k = 1:6
2      mean(grandlinexp(0, 1, 10 ^ k))
3  end
    
```

Dans la suite du problème, on note  $(X_n)_{n \in \mathbb{N}^*}$  une suite de variables aléatoires indépendantes suivant chacune la loi exponentielle linéaire  $\mathcal{E}_\ell(a, b)$  dont les paramètres  $a > 0$  et  $b > 0$  sont inconnus. Soit  $h$  un entier supérieur ou égal à 2. On suit pendant une période de  $h$  années, une « cohorte » de  $n$  individus de même âge au début de l'étude et on modélise leurs durées de vie respectives à partir de cette date par les variables aléatoires  $X_1, X_2, \dots, X_n$ .

**Partie II. Premier décès et intervalle de confiance de  $a$**

Pour tout  $n \in \mathbb{N}^*$ , on définit les variables aléatoires  $M_n, H_n$  et  $U_n$  par :

$$M_n = \min(X_1, X_2, \dots, X_n), \quad H_n = \min(h, X_1, X_2, \dots, X_n) \quad \text{et} \quad U_n = nH_n.$$

- 7. Calculer pour tout  $x \in \mathbb{R}_+$ , la probabilité  $\mathbb{P}([M_n \geq x])$ .  
 Reconnaître la loi de la variable aléatoire  $M_n$ .
- 8. Pour tout  $n \in \mathbb{N}^*$ , on note  $F_{U_n}$  la fonction de répartition de la variable aléatoire  $U_n$ .
  - a) Montrer que pour tout  $x \in \mathbb{R}$ , on a :  $F_{U_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \exp\left(-ax - \frac{b}{2n}x^2\right) & \text{si } 0 \leq x < nh \\ 1 & \text{si } x \geq nh \end{cases}$ .
  - b) Étudier la continuité de la fonction  $F_{U_n}$ .
  - c) La variable aléatoire  $U_n$  admet-elle une densité ?
  - d) Montrer que la suite de variables aléatoires  $(U_n)_{n \in \mathbb{N}^*}$  converge en loi vers une variable aléatoire dont on précisera la loi.
- 9. Soit  $\alpha \in ]0, 1[$ .
  - a) Soit  $Y$  une variable aléatoire qui suit la loi exponentielle de paramètre 1.  
 Trouver deux réels  $c$  et  $d$  strictement positifs tels que :

$$\mathbb{P}([c \leq Y \leq d]) = 1 - \alpha \quad \text{et} \quad \mathbb{P}([Y \leq c]) = \frac{\alpha}{2}$$

- b) Montrer que  $\left[ \frac{c}{U_n}, \frac{d}{U_n} \right]$  est un intervalle de confiance asymptotique de  $a$ , de niveau de confiance  $1 - \alpha$ .

### Partie III. Nombre de survivants et estimateur convergent de $b$

Pour tout  $i \in \mathbb{N}^*$ , soit  $S_i$  et  $D_i$  les variables aléatoires telles que :

$$S_i = \begin{cases} 1 & \text{si } X_i \geq h \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad D_i = \begin{cases} 1 & \text{si } X_i \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Pour tout  $n \in \mathbb{N}^*$ , on pose :  $\bar{S}_n = \frac{1}{n} \sum_{i=1}^n S_i$  et  $\bar{D}_n = \frac{1}{n} \sum_{i=1}^n D_i$ .

10. a) Justifier que pour tout  $i \in \llbracket 1, n \rrbracket$ , on a  $\mathbb{E}(S_i) = G_{a,b}(h)$  et calculer  $\mathbb{E}(S_i D_i)$ .

b) Pour quels couples  $(i, j) \in \llbracket 1, n \rrbracket^2$ , les variables aléatoires  $S_i$  et  $D_j$  sont-elles indépendantes ?

c) Dédire des questions précédentes l'expression de la covariance  $\text{Cov}(\bar{S}_n, \bar{D}_n)$  de  $\bar{S}_n$  et  $\bar{D}_n$  en fonction de  $n$ ,  $G_{a,b}(h)$  et  $G_{a,b}(1)$ . Le signe de cette covariance était-il prévisible ?

11. a) Montrer que  $\bar{S}_n$  est un estimateur sans biais et convergent du paramètre  $G_{a,b}(h)$ .

b) De quel paramètre,  $\bar{D}_n$  est-il un estimateur sans biais et convergent ?

12. On pose :  $z(a, b) = \ln(G_{a,b}(1))$  et  $r(a, b) = \ln(G_{a,b}(h))$ .

Pour tout  $n \in \mathbb{N}^*$ , on pose :  $Z_n = \ln\left(1 - \bar{D}_n + \frac{1}{n}\right)$  et  $R_n = \ln\left(\bar{S}_n + \frac{1}{n}\right)$ .

On admet que  $Z_n$  et  $R_n$  sont des estimateurs convergents de  $z(a, b)$  et  $r(a, b)$  respectivement.

a) Soit  $\varepsilon$ ,  $\lambda$  et  $\mu$  des réels strictement positifs.

(i) Justifier l'inclusion suivante :

$$\left[ |(\lambda Z_n - \mu R_n) - (\lambda z(a, b) - \mu r(a, b))| \geq \varepsilon \right] \subset \left[ \lambda |Z_n - z(a, b)| + \mu |R_n - r(a, b)| \geq \varepsilon \right].$$

(ii) En déduire l'inégalité suivante :

$$\mathbb{P}\left( |(\lambda Z_n - \mu R_n) - (\lambda z(a, b) - \mu r(a, b))| \geq \varepsilon \right) \leq \mathbb{P}\left( \left[ |Z_n - z(a, b)| \geq \frac{\varepsilon}{2\lambda} \right] \right) + \mathbb{P}\left( \left[ |R_n - r(a, b)| \geq \frac{\varepsilon}{2\mu} \right] \right).$$

b) Pour tout  $n \in \mathbb{N}^*$ , on pose :  $B_n = \frac{2}{h-1} Z_n - \frac{2}{h(h-1)} R_n$ .

Montrer que  $B_n$  est un estimateur convergent du paramètre  $b$ .