MATHÉMATIQUES - HEC E 2017

Proposition de corrigé par David Meneu

Lycée Champollion - Grenoble, pour Major-Prépa

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 $\mathcal{B}_n(\mathbb{R})$ 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 $\mathcal{B}_2(\mathbb{R})$ définie par : $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

   a) Le calcul matriciel donne : $A^2 = \begin{pmatrix} 0^2+1^2 & 0\times 1+1\times 0 \\ 0\times 1+1\times 0 & 1^2+0^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$.

   b) Le résultat précédent s'écrit aussi : $A^2 - I_2 = 0_2$, donc que $P(X) = X^2 - 1$ est un polynôme annulateur de A. Les racines (évidentes) de P sont -1 et 1, qui sont donc les seules valeurs propres *possibles* de A.

   Il reste à vérifier que : $A - (-1).I_2 = A + I_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ est non-inversible puisque ses deux lignes sont égales, donc -1 est bien valeur propre de A.

   De même : $A - 1.I_2 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ est non-inversible car ses deux lignes sont opposées, donc proportionnelles : 1 est aussi valeur propre de A. Il n'y a pas d'autres valeurs propres possibles, donc $\mathrm{Sp}(A) = \{-1, 1\}$.

   c) La matrice A est symétrique réelle, donc diagonalisable d'après le théorème du cours. On dispose aussi de l'argument selon lequel c'est une matrice carrée d'ordre 2 qui possède 2 valeurs propres distinctes, et vérifie donc le critère suffisant pour être diagonalisable.

2. *Exemple 2.* Soit B la matrice de $\mathcal{B}_3(\mathbb{R})$ définie par : $B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

   On considère les instructions `Scilab` :

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

   a) D'après l'énoncé, `Scilab` donne : $P^{-1}BP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, B est donc diagonalisable puisqu'elle est semblable à une matrice diagonale, appelons-la D ; les valeurs propres de B sont alors les éléments diagonaux de D, c'est-à-dire :

$$\mathrm{Sp}(B) = \{-1, 1\}$$

   b) La matrice P utilisée par l'énoncé, manifestement inversible, permet d'en déduire immédiatement les sous-espaces propres de B : les colonnes 1 et 3 de P, à savoir $X_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ et $X_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ sont vecteurs propres de B pour la valeur propre 1, et forment une famille libre car ils sont non-colinéaires. Par conséquent : $\dim E_1(B) \geqslant 2$.

   La deuxième colonne de P, à savoir $X_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, est vecteur propre de B pour la valeur propre -1, et $\dim E_{-1}(B) \geqslant 1$.

Il reste à rappeler que, d'après le théorème spectral, $\dim E_1(B) + \dim E_{-1}(B) \leqslant 3$ alors qu'ici :
$\dim E_1(B) + \dim E_{-1}(B) \geqslant 2 + 1 = 3$ pour pouvoir conclure :
$\dim E_1(B) + \dim E_{-1}(B) = 3$ et $\dim E_1(B) = 2$, $\dim E_{-1}(B) = 1$, donc (X_1, X_3) est une base de $E_1(B)$, et (X_2) est une base de $E_{-1}(B)$.

3. a) Par définition, une matrice de $\mathcal{B}_n(\mathbb{R})$ est constituée de n^2 coefficients dont chacun n'a que 2 valeurs possibles : 0 ou 1. Il y a donc $\boxed{2^{n^2}}$ matrices différentes possibles dans $\mathcal{B}_n(\mathbb{R})$.

   b) Pour dénombrer les matrices de $\mathcal{B}_n(\mathbb{R})$ dont chaque ligne et chaque colonne comporte exactement un coefficient égal à 1, on raisonne par choix successifs :

   La ligne 1 d'une telle matrice doit comporter un seul 1, il y a n colonnes possibles pour celui-ci, les autres coefficients étant nuls.
   La ligne 2 doit aussi comporter un seul 1, qui doit cependant ne pas se trouver dans la même colonne que celui de la ligne 1, il n'y a plus que $(n-1)$ possibilités.
   etc... La ligne n doit comporter un seul 1, qui doit être dans une colonne différente des $(n-1)$ colonnes distinctes déjà choisies auparavant : il n'y a plus qu'une seule possibilité !

   Ainsi, par choix successifs, le nombre de ces matrices est : $n \times (n-1) \times \ldots \times 1 = n!$.

   A posteriori, on peut faire le lien entre factorielle est permutations : une matrice du type considéré dans cette question, peut toujours être obtenue d'une seule façon à partir de la matrice identité I_n dont il suffit de *permuter* les lignes de toutes les façons possibles : il y a bien $n!$ permutations possibles de ces n lignes.

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

   On suppose que $u \circ u = id$.

   a) Montrons que $\text{Im}(u - id) \subset F = \text{Ker}(u + id)$, pour cela prenons un élément quelconque x de $\text{Im}(u - id)$: par définition, il existe donc $z \in E$ tel que $x = (u - id)(z)$.
   Mais alors : $(u + id)(x) = (u + id) \circ (u - id)(z) = (u^2 - id^2)(z) = (u^2 - id)(z) = 0_E$ puisque $u^2 = id$, ce qui prouve que $x \in \text{Ker}(u + id) = F$, et donne donc l'inclusion voulue.

   b) En termes de dimensions, l'inclusion $\text{Im}(u - id) \subset \text{Ker}(u + id)$ donne : $\dim \text{Im}(u - id) \leqslant \dim \text{Ker}(u + id)$.
   Or, d'après le théorème du rang : $\dim \text{Im}(u - id) + \dim \text{Ker}(u - id) = \dim E \iff \dim \text{Im}(u - id) = n - q$.

   Par conséquent, l'inégalité précédente se réécrit : $n - q \leqslant p \iff n \leqslant p + q$.

   On suppose désormais que $1 \leqslant p < q$. Soit $(f_1, f_2, \ldots, f_p)$ une base de F et $(g_1, g_2, \ldots, g_q)$ une base de G.

   c) Notons ici que $F = \text{Ker}(u + id)$ et $G = \text{Ker}(u - id)$ sont en fait les deux sous-espaces propres de l'endomorphisme u, respectivement associés aux valeurs propres -1 et 1.

   Une propriété de cours affirme qu'alors, la réunion d'une base de F avec une base de G forme une famille libre, redémontrons-la ici en considérant $p + q$ réels notés $a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q$ tels que :

$$a_1.f_1 + a_2.f_2 + \ldots + a_p.f_p + b_1.g_1 + b_2.g_2 + \ldots + b_q.g_q = 0_E \qquad (\star)$$

   Puisque les vecteurs f_i sont tous éléments de $F = E_1(u)$, ils vérifient : $u(f_i) = -f_i$ pour tout i compris entre 1 et p. De même, pour tout i compris entre 1 et q, $u(g_i) = g_i$ puisque $g_i \in G = E_{-1}(u)$.
   La composition par u de la relation $(\star)$, et la linéarité de l'endomorphisme donnent :

$$a_1.u(f_1) + a_2.u(f_2) + \ldots + a_p.u(f_p) + b_1.u(g_1) + b_2.u(g_2) + \ldots + b_q.u(g_q) = u(0_E)$$

$$\iff -a_1.f_1 - a_2.f_2 - \ldots - a_p.f_p + b_1.g_1 + b_2.g_2 + \ldots + b_q.g_q = 0_E \qquad (\star\star)$$

   La somme membre à membre des égalités $(\star)$ et $(\star\star)$ donne alors :
   $2b_1.g_1 + 2b_2.g_2 + \ldots + 2b_q.g_q = 0_E$ qui entraîne : $b_1 = b_2 = \ldots = b_q = 0$ puisque $(g_1, g_2, \ldots, g_q)$ est une famille libre.

La relation $(\star)$ devient alors : $a_1.f_1 + a_2.f_2 + \ldots + a_p.f_p = 0_E \Longrightarrow a_1 = a_2 = \ldots = a_p = 0$ puisque $(f_1, f_2, \ldots, f_p)$ est aussi libre.

Bref, on a prouvé :

$$a_1.f_1+a_2.f_2+\ldots+a_p.f_p+b_1.g_1+b_2.g_2+\ldots+b_q.g_q = 0_E \Longrightarrow a_1 = a_2 = \ldots = a_p = b_1 = b_2 = \ldots = b_q = 0$$

ce qui prouve bien que la famille $(f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q)$ est une famille libre de $p+q$ vecteurs de l'espace vectoriel E qui est de dimension n : comme $p+q \geqslant n$ d'après b), cela implique nécessairement que $p+q = n$ car le cardinal d'une famille libre ne peut pas excéder la dimension de l'espace : on dispose ainsi d'une famille libre de n vecteurs de E, c'est bien une base de cet espace vectoriel.

d) Par linéarité de u, et d'après les relations vérifiées par les vecteurs f_i et g_i rappelés ci-dessus :

$$u(g_1 - f_1) = u(g_1) - u(f_1) = g_1 + f_1 \quad \text{et} \quad u(g_1 + f_1) = u(g_1) + u(f_1) = g_1 - f_1$$

e) Les calculs de la question précédente permettent de voir que l'endomorphisme u a pour action d'envoyer le vecteur $g_1 - f_1$ sur le vecteur $g_1 + f_1$, et réciproquement : cela doit nous inciter à considérer toutes les paires de vecteurs $g_i - f_i$ et $g_i + f_i$, où i peut varier de 1 à p puisqu'on a supposé $p < q$ (il y a plus de vecteurs g_i que de vecteurs f_i).

Considérons donc la famille $\mathcal{C} = (g_1 - f_1, g_1 + f_1, g_2 - f_2, g_2 + f_2, \ldots, g_p - f_p, g_p + f_p, g_{p+1}, \ldots, g_q)$: elle rassemble tous les couples évoqués plus haut, les $p - q$ vecteurs restants de la base de G ne pouvant être appariés avec un vecteur de la base de F, étant repris seuls.

C'est une famille de $2p+(q-p) = p+q = n$ vecteurs de E, dont on vérifie qu'elle est libre en considérant n réels $(a_1, b_1, a_2, b_2, \ldots, a_p, b_p, c_{p+1}, \ldots, c_q)$ tels que :

$$\begin{aligned} & a_1.(g_1 - f_1) + b_1.(g_1 + f_1) + a_2.(g_2 - f_2) + b_2.(g_2 + f_2) + \ldots + a_p.(g_p - f_p) + b_p.(g_p + f_p) \\ & \qquad\qquad + c_{p+1}.g_{p+1} + \ldots + c_q.g_q = 0_E \\ \Longleftrightarrow\ & (-a_1 + b_1).f_1 + (a_1 + b_1).g_1 + (-a_2 + b_2).f_2 + (a_2 + b_2).g_2 + \ldots + (-a_p + b_p).f_p + (a_p + b_p).g_p \\ & \qquad\qquad + c_{p+1}.g_{p+1} + \ldots + c_q.g_q = 0_E \\ \Longleftrightarrow\ & \begin{cases} b_1 - a_1 = 0 \\ b_1 + a_1 = 0 \end{cases} \text{ et } \begin{cases} b_2 - a_2 = 0 \\ b_2 + a_2 = 0 \end{cases} \text{ et } \cdots \text{ et } \begin{cases} b_p - a_p = 0 \\ b_p + a_p = 0 \end{cases} \text{ et } b_{p+1} = \cdots = b_q = 0 \end{aligned}$$

puisque $(f_1, f_2, \ldots, f_p, g_1, g_2, \ldots, g_q)$ est une base de E. Chaque système $\begin{cases} b_i - a_i = 0 \\ b_i + a_i = 0 \end{cases}$ a pour unique solution $a_i = b_i = 0$ (réaliser l'opération $L_2 \leftarrow L_1 + L_2$ pour s'en convaincre), dont finalement :

$$\begin{aligned} & a_1.(g_1 - f_1) + b_1.(g_1 + f_1) + a_2.(g_2 - f_2) + b_2.(g_2 + f_2) + \ldots + a_p.(g_p - f_p) + b_p.(g_p + f_p) \\ & \qquad\qquad + c_{p+1}.g_{p+1} + \ldots + c_q.g_q = 0_E \\ \Longrightarrow\ & a_1 = b_1 = a_2 = b_2 = \ldots = a_p = b_p = c_{p+1} = \ldots = c_q = 0 \end{aligned}$$

et la famille $\mathcal{C}$ est bien libre, c'est donc une case de E puisque $\text{Card}(\mathcal{C}) = n = \dim E$.

Dans cette base : $\forall i \in [\![1, p,]\!]$, $u(g_i - f_i) = g_i + f_i$ et $u(g_i + f_i) = g_i - f_i$, tandis que : $\forall i \in [\![p+1, q]\!]$, $u(g_i) = g_i$, donc la matrice de u dans la base $\mathcal{C}$ est de la forme :

$$\text{Mat}_{\mathcal{C}}(u) = \begin{pmatrix} 0 & 1 & & & & & & & & \\ 1 & 0 & & & & & & (0) & & \\ & & 0 & 1 & & & & & & \\ & & 1 & 0 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 0 & 1 & & & \\ & & & & & 1 & 0 & & & \\ & & & & & & & 1 & & \\ & (0) & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{pmatrix}$$

On commence par p "blocs diagonaux" de même forme que la matrice A de la toute première question, suivi d'une partie diagonale de $q - p$ coefficients tous égaux à 1.

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

Partie I. Loi exponentielle linéaire

1. a) La fonction $G_{a,b}$ est la composée de la fonction trinôme $x \mapsto -ax - \dfrac{b}{2}x^2$ qui est clairement strictement décroissante sur $\mathbb{R}_+$ (sa dérivée est $x \mapsto -a - bx$, strictement négative sur $\mathbb{R}^+$ car a et b sont strictement positifs), et de la fonction exp, strictement croissante sur $\mathbb{R}$: $G_{a,b}$ est ainsi strictement décroissante sur $\mathbb{R}_+$ par composition. Elle est aussi continue sur $\mathbb{R}_+$ comme composée de fonctions continues, et :

   $G_{a,b}(0) = e^0 = 1$, tandis que $\lim\limits_{x\to+\infty} -ax - \dfrac{b}{2}x^2 - \infty$, donc $\lim\limits_{x\to+\infty} G_{a,b}(x) = \lim\limits_{X\to-\infty} e^X = 0$: d'après le théorème éponyme, $G_{a,b}$ réalise bien une bijection de $\mathbb{R}$ dans $]0,1]$.

   b) Pour $y > 0$, on considère l'équation d'inconnue $x \in \mathbb{R}$: $\;ax + \dfrac{b}{2}x^2 = y \iff bx^2 + 2ax - 2y = 0$.

   C'est une équation du second degré, de discriminant : $\Delta = 4a^2 + 8by > 0$ puisque a, b, y sont strictement positifs. L'équation admet donc deux solutions distinctes :

$$x_1 = \frac{-2a - \sqrt{4a^2 + 8by}}{2b} = \frac{-a - \sqrt{a^2 + 2by}}{b} \quad \text{et} \quad x_2 = \frac{-a + \sqrt{a^2 + 2by}}{b}$$

   c) Pour trouver l'expression de $G_{a,b}^{-1}$, bijection réciproque de $G_{a,b}$, on doit trouver pour tout $y \in]0,1]$, l'unique solution $x \in \mathbb{R}_+$ de l'équation $G_{a,b}(x) = y$:

$$G_{a,b}(x) = y \iff \exp\Big(-ax - \frac{b}{2}x^2\Big) = y \iff -ax - \frac{b}{2}x^2 = \ln(y) \iff ax + \frac{b}{2}x^2 = -\ln(y)$$

   Comme $y \in]0,1]$, $\;\ln(y) \leqslant 0 \iff -\ln(y) \geqslant 0$, donc la résolution de l'équation faite en b) fournit deux solutions : $\quad x_1 = \dfrac{-a - \sqrt{a^2 - 2b\ln(y)}}{b}$ et $x_2 = \dfrac{-a + \sqrt{a^2 - 2b\ln(y)}}{b}$.

   Il est évident que la première solution est négative, donc ne peut convenir. Par contre :
   comme $\ln(y) \leqslant 0$, $\;a^2 - 2b\ln(y) \geqslant a^2$, donc $\sqrt{a^2 - 2b\ln(y)} \geqslant a$ par stricte croissance de la fonction racine carrée sur $\mathbb{R}_+$, et x_2 est une solution positive.
   Notons aussi que dans le cas où $y = 1$, les deux solutions sont confondues, égales à 0 ; bref :

$$\forall y \in]0,1], \quad G_{a,b}^{-1}(y) = \frac{-a + \sqrt{a - 2b\ln(y)}}{b}$$

   Et enfin : pour tout $u \in [0,1[$, $\;1 - u \in]0,1]$ donc : $\quad G_{a,b}^{-1}(1-u) = \dfrac{-a + \sqrt{a^2 - 2b\ln(1-u)}}{b}$.

2. a) Remarquons ici que : $\forall x \in \mathbb{R}_+$, $\;-ax - \dfrac{b}{2}x^2 \leqslant -ax$, donc $\exp\Big(-ax - \dfrac{b}{2}x^2\Big) = G_{a,b}(x) \leqslant e^{-ax}$ par croissance de exp sur $\mathbb{R}$: or $\displaystyle\int_0^{+\infty} e^{-ax}\mathrm{d}x$ est une intégrale convergente, qui est à un facteur a près l'intégrale d'une densité de loi exponentielle de paramètre a, et vaut $\dfrac{1}{a}$.
   Les fonctions concernées sont continues, positives sur $\mathbb{R}_+$, donc d'après le théorème de comparaison d'intégrales de telles fonctions, $\displaystyle\int_0^{+\infty} G_{a,b}(x)\mathrm{d}x$ est une intégrale convergente.

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

Une densité de la loi normale de paramètres (m, σ^2) est la fonction $x \mapsto \dfrac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{(x-m)^2}{2\sigma^2}}$, ce qui correspond bien à la fonction f avec :

$$m = -\frac{a}{b} \text{ (espérance)} \quad \text{et} \quad \sigma^2 = \frac{1}{b} \text{ (variance)}$$

c) Soit Φ la fonction de répartition de la loi normale centrée réduite :

Pour tout réel $x \in \mathbb{R}_+$, $-\dfrac{1}{2}b(x+\dfrac{a}{b})^2 = -\dfrac{1}{2}b(x^2 + \dfrac{2a}{b}x + \dfrac{a^2}{b^2}) = -ax - \dfrac{b}{2}x^2 - \dfrac{a^2}{2b}$, donc :

$\forall x \in \mathbb{R}_+,\ G_{a,b}(x) = \exp\big(-ax - \dfrac{b}{2}x^2\big) = \exp\Big(-\dfrac{1}{2}b(x+\dfrac{a}{b})^2\Big) \times \exp\big(\dfrac{a^2}{2b}\big) = \sqrt{\dfrac{2\pi}{b}} \times \exp\big(\dfrac{a^2}{2b}\big) \times f(x)$, donc :

$$\begin{aligned}\int_0^{+\infty} G_{a,b}(x)\mathrm{d}x &= \sqrt{\frac{2\pi}{b}} \times \exp\big(\frac{a^2}{2b}\big) \times \int_0^{+\infty} f(x)\mathrm{d}x \\ &= \sqrt{\frac{2\pi}{b}} \times \exp\big(\frac{a^2}{2b}\big) \times P(Y \geqslant 0) \qquad \text{avec } Y \text{ suivant la loi normale } \mathcal{N}(-\frac{a}{b}, \frac{1}{b})\end{aligned}$$

Or on sait, d'après le cours sur la loi normale, que $Y^* = \dfrac{Y - E(Y)}{\sigma(Y)} = \dfrac{Y + \frac{a}{b}}{\frac{1}{\sqrt{b}}} = \sqrt{b}Y + \dfrac{a}{\sqrt{b}}$ suit la loi normale centrée, réduite, et :

$P(Y \geqslant 0) = P\Big(Y^* \geqslant \dfrac{a}{\sqrt{b}}\Big) = 1 - \Phi\Big(\dfrac{a}{\sqrt{b}}\Big) = \Phi\Big(-\dfrac{a}{\sqrt{b}}\Big)$ d'après les propriétés de la loi normale, donc :

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

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

a) La fonction $f_{a,b}$ est constante nulle sur $]-\infty; 0[$, donc continue, positive sur cet intervalle.
La fonction $f_{a,b}$ est continue, positive sur $]0; +\infty[$ comme produit de deux telles fonctions : $f_{a,b}$ est donc continue, positive sur $\mathbb{R}$ sauf peut-être en 0. De plus, sous réserve de convergence :
$\displaystyle\int_{-\infty}^{+\infty} f_{a,b}(x)\mathrm{d}x = \int_{-\infty}^{0} f_{a,b}(x)\mathrm{d}x + \int_0^{+\infty} f_{a,b}(x)\mathrm{d}x = \int_0^{+\infty} f_{a,b}(x)\mathrm{d}x$; pour tout réel $A > 0$:

$$\begin{aligned}\int_0^A f_{a,b}(x)\mathrm{d}x = \int_0^A (a+bc)\exp\Big(-ax - \frac{b}{2}x^2\Big)\mathrm{d}x &= \Big[-\exp\Big(-ax - \frac{b}{2}x^2\Big)\Big]_0^A \\ &= -\exp\Big(-a.A - \frac{b}{2}.A^2\Big) + 1 \xrightarrow[A \to +\infty]{} 1\end{aligned}$$

ce qui prouve que $\displaystyle\int_0^{+\infty} f_{a,b}(x)\mathrm{d}x = 1 = \int_{-\infty}^{+\infty} f_{a,b}(x)\mathrm{d}x$, et donc que $f_{a,b}$ est bien 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 $\mathcal{E}_\ell(a,b)$. La variable aléatoire X admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x f_{a,b}(x)\mathrm{d}x$ est absolument convergente.

Comme la fonction $x \mapsto x f_{a,b}(x)$ est nulle sur $]-\infty, 0[$ et positive sur $[0, +\infty[$, cela revient à prouver la convergence simple de l'intégrale $\displaystyle\int_0^{+\infty} x f_{a,b}(x)\mathrm{d}x = \int_0^{+\infty} x(a+bx)\exp\Big(-ax - \frac{b}{2}x^2\Big)$.

Soit $A > 0$, on pose :

$$\begin{aligned} u(x) &= x & \longrightarrow \quad & u'(x) = 1 \\ v'(x) &= (a+bx)\exp\Big(-ax-\frac{b}{2}x^2\Big) & \longrightarrow \quad & v(x) = -\exp\Big(-ax-\frac{b}{2}x^2\Big) = -G_{a,b}(x) \end{aligned}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}_+$, donc par intégration par parties :

$$\int_0^A x f_{a,b}(x)\mathrm{d}x = \Big[-xG_{a,b}(x)\Big]_0^A + \int_0^A G_{a,b}(x)\mathrm{d}x = -A\exp\Big(-a.A-\frac{b}{2}.A^2\Big) + \int_0^A G_{a,b}(x)\mathrm{d}x$$

Par croissances comparées : $\lim\limits_{A\to+\infty} -Ae^{-a.A} = 0$, donc $\lim\limits_{A\to+\infty} -A.e^{-a.A}\times\exp\Big(-\frac{b}{2}.A^2\Big) = 0\times 0 = 0$.

Comme l'intégrale $\displaystyle\int_0^{+\infty} G_{a,b}(x)\mathrm{d}x$ est convergente d'après la question 2.a), l'intégrale $\displaystyle\int_0^{+\infty} xf_{a,b}(x)\mathrm{d}x$ est elle-même convergente, et X admet une espérance qui s'écrit :

$$E(X) = \int_0^{+\infty} G_{a,b}(x)\mathrm{d}x$$

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

   a) Pour tout réel $x \in \mathbb{R}_+$:

$$\begin{aligned} P(X\geqslant x) &= P\Big(\frac{-a+\sqrt{a^2+2bY}}{b}\geqslant x\Big) = P(-a+\sqrt{a^2+2bY}\geqslant bx) \\ &= P(\sqrt{a^2+2bY}\geqslant a+bx) = P(a^2+2bY\geqslant(a+bx)^2) \quad \text{par stricte croissante de } t\mapsto t^2 \text{ sur } \mathbb{R}_+ \\ &= P(2bY\geqslant 2abx+b^2x^2) = P\Big(Y\geqslant ax+\frac{b}{2}x^2\Big) \end{aligned}$$

   Au vu de la loi de Y, et vu que $ax+\frac{b}{2}x^2\geqslant 0$:

$$\forall x\in\mathbb{R}_+, \quad P(X\geqslant x) = P\Big(Y\geqslant ax+\frac{b}{2}x^2\Big) = \exp\Big(-ax-\frac{b}{2}x^2\Big) = G_{a,b}(x)$$

   b) Le calcul précédent permet d'en déduire la fonction de répartition de X :
   pour tout $x\in\mathbb{R}_+$, $P(X<x) = 1-P(X\geqslant x) = 1-G_{a,b}(x)$. Il est également clair que par définition, et comme on l'a vu à la question 1.b), X est positive, donc : $\forall x\in\mathbb{R}_-$, $P(X<x)=0$.

   La fonction $x\mapsto P(X<x)$ est ainsi de classe $\mathcal{C}^1$, donc aussi continue sur $]-\infty,0[$ et $]0,+\infty[$. Et comme : $\lim\limits_{x\to0^+} 1-G_{a,b}(x) = 1-e^0 = 0 = \lim\limits_{x\to0^-} P(X<x)$, la fonction $x\mapsto P(X<x)$ est continue sur $\mathbb{R}$ tout entier.

   On en déduit que : $\forall x\in\mathbb{R}$, $F_X(x) = P(X\leqslant x) = P(X<x)$ et que X est une variable à densité. Une densité de X est obtenue par dérivation de F_X sur $\mathbb{R}$, sauf en 0 :

$$\forall x\in\mathbb{R}^*, \quad F_X'(x) = \begin{cases} -(-a-bx)\exp\Big(-ax-\frac{b}{2}x^2\Big) = (a+bx)\exp\Big(-ax-\frac{b}{2}x^2\Big) & \text{si } x>0 \\ 0 & \text{si } x<0 \end{cases}$$

   la fonction $f_{a,b}$ est donc une densité de X, qui suit bien la loi $\mathcal{E}_\ell(a,b)$.

   c) On note U une variable aléatoire suivant la loi uniforme sur $[0,1[$: la variable $1-U$ est alors à valeurs dans $]0,1]$, donc la variable aléatoire $G_{a,b}^{-1}(1-U)$ est bien définie, à valeurs dans $\mathbb{R}_+$, donc :
   $\forall x\in\mathbb{R}_-$, $P(G_{a,b}^{-1}(1-U)\leqslant x) = 0$, et :

$$\begin{aligned} \forall x\in\mathbb{R}_+, \quad P(G_{a,b}^{-1}(1-U)\leqslant x) &= P(1-U\geqslant G_{a,b}(x)) \quad \text{car } G_{a,b} \text{ est bijective, strictement décroissante sur } \mathbb{R}_+ \\ &= P(U\leqslant 1-G_{a,b}(x)) \end{aligned}$$

   or : $\forall x\in\mathbb{R}_+$, $G_{a,b}(x)\in]0,1] \iff 1-G_{a,b}(x)\in[0;1[$ donc :

$$\forall x\in\mathbb{R}_+, \quad P(G_{a,b}^{-1}(1-U)\leqslant x) = 1-G_{a,b}(x)) = F_X(x)$$

   Donc $G_{a,b}^{-1}(1-U)$ suit la même loi $\mathcal{E}_\ell(a,b)$ que X.

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

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

a) La ligne de code (2) crée un vecteur aléatoire de n simulations indépendantes de la loi uniforme à densité sur $[0,1[$.

b) Selon la méthode d'inversion habituelle, l'instruction en ligne 3 doit être `y = -log(1-u)` qui est une opération terme à terme sur tous les éléments du vecteur `u`, telle que le vecteur `y` contient alors n simulations indépendantes de la loi exponentielle de paramètre 1.

De la sorte, le vecteur `x` contient bien, d'après la question 4., n simulations indépendantes de la loi exponentielle linéaire $\mathcal{E}_\ell(a,b)$.

6. La boucle `Scilab` suivante :

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

calcule la moyenne empirique $\overline{X}_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ d'un échantillon $(X_1,\dots,X_n)$ de n simulations de la loi $\mathcal{E}_\ell(a,b)$, pour une taille $n=10^k$ d'échantillon dix fois plus importante à chaque fois : $n=10$, puis 100, puis 1000,10000, 1000000. La loi faible des grands nombres assure que la suite $(\overline{X}_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers la constante $E(X)$, qui est ainsi estiméede plus en plus précisément par la moyenne empirique d'un échantillon de plus en plus important.

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. Soit $x\in\mathbb{R}_+$, par définition du minimum d'un échantillon :

$$[M_n \geqslant x] = \bigcap_{i=1}^{n}[X_i \geqslant x]$$

Ainsi, par indépendance des variables aléatoires X_i :

$$\begin{aligned}\forall x\in\mathbb{R}_+,\quad P(M_n\geqslant x) &= \prod_{i=1}^{n} P(X_i\geqslant x)\\ &= \Big(P(X\geqslant x)\Big)^n \qquad \text{car les } X_i \text{ suivent toutes la même loi que } X\\ &= (G_{a,b}(x))^n = \exp\Big(-anx-\frac{bn}{2}x^2\Big) = G_{na,nb}(x)\end{aligned}$$

Cela suffit, d'après la question 4.b), pour en déduire que M_n suit la loi exponentielle linéaire $\mathcal{E}_\ell(na,nb)$.

8. Pour tout $n\in\mathbb{N}^*$, on note F_{U_n} la fonction de répartition de la variable aléatoire U_n.

a) Comme h est positif, et que les variables $(X_i)_{i\in\mathbb{N}^*}$ sont à valeurs positives, $U_n = nH_n = \min(h, X_1, X_2, \dots, X_n)$ est à valeurs positives, donc : $\forall x < 0,\ F_{U_n}(x) = P(U_n \leqslant x) = 0$.

Pour tout réel positif $x \in \mathbb{R}_+$: $F_{U_n}(x) = P(U_n \leqslant x) = P\big(H_n \leqslant \frac{x}{n}\big) = P\big(\min(h, X_1, X_2, \dots, X_n) \leqslant \frac{x}{n}\big)$.

Si $h \leqslant \frac{x}{n} \iff nh \leqslant x$: puisque $\min(h, X_1, X_2, \dots, X_n) \leqslant h$, alors l'événement $\big[\min(h, X_1, X_2, \dots, X_n) \leqslant \frac{x}{n}\big]$ est certain et $F_{U_n}(x) = 1$.

Si $0 \leqslant x < nh \iff 0 \leqslant \frac{x}{n} < h$: $\big[\min(h, X_1, X_2, \dots, X_n) \leqslant \frac{x}{n}\big]$ est réalisé si et seulement si $\big[\min(X_1, X_2, \dots, X_n) \leqslant \frac{x}{n}\big] = \big[M_n \leqslant \frac{x}{n}\big]$ l'est (h ne peut pas être le minimum pour que cet événement soit réalisé), et :

$$F_{U_n}(x) = P\big(M_n \leqslant \frac{x}{n}\big) = 1 - \exp\Big(-na.\frac{x}{n} - \frac{nb}{2}.(\frac{x}{n})^2\Big) = 1 - \exp\Big(-ax - \frac{b}{2n}x^2\Big)$$

Bilan : pour tout $x \in \mathbb{R},\ F_{U_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \exp\Big(-ax - \dfrac{b}{2n}x^2\Big) & \text{si } 0 \leqslant x < nh \\ 1 & \text{si } x \geqslant nh \end{cases}$.

b) La fonction F_{U_n} est continue sur $]-\infty, 0[$ et $]nh, +\infty[$ comme fonction constante, et sur $]0, nh[$ comme somme et composée de fonctions continues sur $\mathbb{R}$. Par ailleurs :

$\lim\limits_{x\to 0^-} F_{U_n}(x) = 0$ et $\lim\limits_{x\to 0^+} F_{U_n}(x) = 1 - e^0 = 0 = F_{U_n}(0)$, donc F_{U_n} est continue en 0.

$\lim\limits_{x\to nh^+} F_{U_n}(x) = 1 = F_{U_n}(nh)$ et $\lim\limits_{x\to nh^-} F_{U_n}(x) = 1 - \exp\big(-anh - \frac{b}{2n}(nh)^2\big) = 1 - \exp\big(-anh - \frac{bnh^2}{2}\big) < 1$, donc F_{U_n} n'est pas continue en nh.

Finalement, F_{U_n} est continue sur $\mathbb{R}$ sauf en nh.

c) Puisque F_{U_n} n'est pas continue sur $\mathbb{R}$, la variable aléatoire U_n n'admet pas de densité.

d) On calcule ici, pour tout réel x, la limite $\lim\limits_{n\to+\infty} F_{U_n}(x)$. On remarque que :

- Pour tout $x < 0$: $\lim\limits_{n\to+\infty} F_{U_n}(x) = \lim\limits_{n\to+\infty} 0 = 0$.
- Pour tout réel positif $x \in \mathbb{R}_+$: puisque $h > 0$, $\lim\limits_{n\to+\infty} nh = +\infty$ et on peut toujours trouver n assez grand pour que $0 \leqslant x < nh$.

  Ainsi : $\forall x \in \mathbb{R}_+,\quad \lim\limits_{n\to+\infty} F_{U_n}(x) = \lim\limits_{n\to+\infty} \exp\Big(-ax - \frac{b}{2n}x^2\Big) = 1 - e^{-ax}$.

  Bilan : $\quad \forall x \in \mathbb{R},\quad \lim\limits_{n\to+\infty} F_{U_n}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-ax} & \text{si } x \geqslant 0 \end{cases}$.

  On en déduit que la suite de variables aléatoires $(U_n)_{n\in\mathbb{N}^*}$ **converge en loi vers la loi exponentielle de paramètre** a.

9. Soit $\alpha \in]0, 1[$.

a) Soit Y une variable aléatoire qui suit la loi exponentielle de paramètre 1 ; on cherche deux réels c et d strictement positifs tels que :

$$\begin{cases} P(c \leqslant Y \leqslant d) = 1 - \alpha \\ P(Y \leqslant c) = \dfrac{\alpha}{2} \end{cases} \iff \begin{cases} P(Y \leqslant d) - P(Y < c) = 1 - \alpha \\ P(Y \leqslant c) = \dfrac{\alpha}{2} \end{cases} \iff \begin{cases} 1 - e^{-d} - (1 - e^{-c}) = 1 - \alpha \\ 1 - e^{-c} = \dfrac{\alpha}{2} \end{cases}$$

$$\iff \begin{cases} e^{-c} - e^{-d} = 1 - \alpha \\ e^{-c} = 1 - \dfrac{\alpha}{2} \end{cases} \iff \begin{cases} e^{-d} = e^{-c} - 1 + \alpha = \dfrac{\alpha}{2} \\ e^{-c} = 1 - \dfrac{\alpha}{2} \end{cases} \iff \begin{cases} d = -\ln(\alpha/2) \\ c = -\ln(1 - \alpha/2) \end{cases}$$

b) Pour démontrer le résultat demandé, on évalue :

$P\Big(\frac{c}{U_n} \leqslant a \leqslant \frac{d}{U_n}\Big) = P(c \leqslant a.U_n \leqslant d) = P\big(\frac{c}{a} \leqslant U_n \leqslant \frac{d}{a}\big) = F_{U_n}\big(\frac{d}{a}\big) - F_{U_n}\big(\frac{c}{a}\big)$, où :

D'après 8.d) : $\lim\limits_{n\to+\infty} F_{U_n}\big(\frac{d}{a}\big) = 1 - e^{-a.\frac{d}{a}} = 1 - e^{-d} = 1 - \frac{\alpha}{2}$, et $\lim\limits_{n\to+\infty} F_{U_n}\big(\frac{c}{a}\big) = 1 - e^{-a.\frac{c}{a}} = 1 - e^{-c} = \frac{\alpha}{2}$, donc :

$$\lim_{n\to+\infty} P\Big(\frac{c}{U_n} \leqslant a \leqslant \frac{d}{U_n}\Big) = \lim_{n\to+\infty} F_{U_n}\big(\frac{d}{a}\big) - F_{U_n}\big(\frac{c}{a}\big) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$$

ce qui prouve bien 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 $i \in \mathbb{N}^*$, soient S_i et D_i les variables aléatoires telles que : $S_i = \begin{cases} 1 & \text{si } X_i \geqslant h \\ 0 & \text{sinon} \end{cases}$ et $D_i = \begin{cases} 1 & \text{si } X_i \leqslant 1 \\ 0 & \text{sinon`} \end{cases}$.

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

10. a) La variable aléatoire S_i est une variable de Bernoulli, donc : $E(S_i) = P(S_i = 1) = P(X_i \geqslant h) = G_{a,b}(h)$.

La variable aléatoire S_iD_i est elle aussi une variable de Bernoulli comme produit de deux variables de Bernoulli, et :

$E(S_iD_i) = P(S_iD_i = 1) = P([S_i = 1] \cap [D_i = 1]) = P([X_i \geqslant h] \cap [X_i \leqslant 1]) = 0$ car puisque $h \geqslant 2$, les événements $[X_i \geqslant h]$ et $[X_i \leqslant 1]$ sont incompatibles.

b) D'après ce qui précède :

$P([S_i = 1] \cap [D_i = 1]) = 0$ alors que $P(S_i = 1) \times P(D_i = 1) = P(X_i \geqslant h) \times P(X_i \leqslant 1) > 0$,

donc pour $i = j$, les variables aléatoires S_i et D_j ne sont pas indépendantes.

Pour $i \neq j$:

$P([S_i = 1] \cap [D_j = 1]) = P([X_i \geqslant h] \cap [X_j \leqslant 1]) = P(X_i \geqslant h) \times P(X_j \leqslant 1) = P(S_i = 1) \times P(D_j = 1)$ car X_i et X_j sont indépendantes pour $i \neq j$.

Cela suffit pour prouver que les variables de Bernoulli S_i et D_j sont indépendantes lorsque $i \neq j$.

c) Pour tout entier $n \in \mathbb{N}^*$, par bilinéarité de la covariance :

$$\begin{aligned}
\operatorname{Cov}(\overline{S}_n, \overline{D}_n) &= \frac{1}{n^2}\operatorname{Cov}\Big(\sum_{i=1}^{n} S_i, \sum_{j=1}^{n} D_j\Big) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\operatorname{Cov}(S_i, D_j) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Cov}(S_i, D_i) \quad \text{car } \operatorname{Cov}(S_i, D_j) = 0 \text{ dès que } i \neq j \\
&= \frac{1}{n^2}\sum_{i=1}^{n} E(S_iD_i) - E(S_i)E(D_i) \quad \text{or } E(S_iD_i) = 0 \\
&= -\frac{1}{n^2}\sum_{i=1}^{n} G_{a,b}(h) \times (1 - G_{a,b}(1)) \\
\operatorname{Cov}(\overline{S}_n, \overline{D}_n) &= -\frac{1}{n}G_{a,b}(h)(1 - G_{a,b}(1))
\end{aligned}$$

La négativité de cette covariance est prévisible : $\overline{S}_n$ est le nombre moyen des variables de l'échantillon $(X_1, \ldots, X_n)$ qui sont supérieures à h (avec $h \geqslant 2$), tandis que $\overline{D}_n$ est le nombre moyen des variables de ce même échantillon qui sont inférieures à 1. Toute augmentation de l'une des deux moyennes, provoque une diminution de l'autre : les deux variables évoluent en sens contraire l'une de l'autre.

11. a) Un résultat classique de cours, est que la moyenne empirique $\frac{1}{n}\sum_{i=1}^{n} S_i$ est un estimateur sans biais et convergent de l'espérance commune $E(S_i) = G_{a,b}(h)$ de l'échantillon $(S_1, S_2, \ldots, S_n)$.

b) La variable aléatoire $\overline{D}_n$, moyenne empirique de l'échantillon $(D_1, D_2, \ldots, D_n)$, est de son côté un estimateur sans biais et convergent de l'espérance $E(D_i) = P(X_i \leqslant 1) = 1 - G_{a,b}(1) = 1 - e^{-a-b/2}$.

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

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

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

a) Soient ε, λ et μ des réels strictement positifs.

(i) D'après l'inégalité triangulaire :

$$\big|(\lambda Z_n-\mu R_n)-(\lambda z(a,b)-\mu r(a,b))\big| = \big|\lambda(Z_n-z(a,b))-\mu(R_n-r(a,b))\big| \leqslant \lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)|$$

donc par transitivité de l'inégalité :
$\big|(\lambda Z_n-\mu R_n)-(\lambda z(a,b)-\mu r(a,b))\big| \geqslant \varepsilon$ implique $\lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)| \geqslant \varepsilon$,
ce qui se traduit bien par l'inclusion d'événements :

$$\Big[\big|(\lambda Z_n-\mu R_n)-(\lambda z(a,b)-\mu r(a,b))\big| \geqslant \varepsilon\Big] \subset \Big[\lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)| \geqslant \varepsilon\Big]$$

(ii) L'inclusion précédente donne, par croissance de la probabilité :

$$P\Big(\Big[\big|(\lambda Z_n-\mu R_n)-(\lambda z(a,b)-\mu r(a,b))\big| \geqslant \varepsilon\Big]\Big) \leqslant P\Big(\Big[\lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)| \geqslant \varepsilon\Big]\Big)$$

Enfin : si la somme $\lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)|$ est supérieure ou égale à ε, alors l'un des deux termes au moins est forcément supérieur à $\dfrac{\varepsilon}{2}$ (sinon par contraposée : $\lambda|Z_n-z(a,b)| < \dfrac{\varepsilon}{2}$ **et** $\mu|R_n-r(a,b)| < \dfrac{\varepsilon}{2}$ impliquent $\lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)| < \varepsilon$), c 'est-à-dire :

$$\Big[\lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)| \geqslant \varepsilon\Big] \subset \Big[\lambda|Z_n-z(a,b)| \geqslant \frac{\varepsilon}{2}\Big] \cup \Big[\mu|R_n-r(a,b)| \geqslant \frac{\varepsilon}{2}\Big]$$

$$\iff \Big[\lambda|Z_n-z(a,b)|+\mu|R_n-r(a,b)| \geqslant \varepsilon\Big] \subset \Big[|Z_n-z(a,b)| \geqslant \frac{\varepsilon}{2\lambda}\Big] \cup \Big[|R_n-r(a,b)| \geqslant \frac{\varepsilon}{2\mu}\Big]$$

Cette inclusion, et la propriété générale : $P(A\cup B) = P(A)+P(B)-P(A\cap B) \leqslant P(A)+P(B)$, valable pour tous événements A et B d'un même espace probabilisé, donnent finalement par transitivité :

$$0 \leqslant P\Big(\Big[\big|(\lambda Z_n-\mu R_n)-(\lambda z(a,b)-\mu r(a,b))\big| \geqslant \varepsilon\Big]\Big) \leqslant P\Big(\Big[|Z_n-z(a,b)| \geqslant \frac{\varepsilon}{2\lambda}\Big]\Big)+P\Big(\Big[|R_n-r(a,b)| \geqslant \frac{\varepsilon}{2\mu}\Big]\Big)$$

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

Le résultat admis par l'énoncé, selon lequel Z_n et R_n sont des estimateurs convergents de $z(a,b)$ et $r(a,b)$ respectivement, implique :

$$\lim_{n\to+\infty} P\Big(\Big[|Z_n-z(a,b)| \geqslant \frac{\varepsilon}{2\lambda}\Big]\Big) = \lim_{n\to+\infty} P\Big(\Big[|R_n-r(a,b)| \geqslant \frac{\varepsilon}{2\mu}\Big]\Big) = 0$$

On n'a pas oublié à la question précédente, d'écrire explicitement la positivité des probabilités concernées par l'inégalité finale, car c'est le **théorème d'encadrement** qui permet de conclure que :

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} P\Big(\Big[\big|(\lambda Z_n-\mu R_n)-(\lambda z(a,b)-\mu r(a,b))\big| \geqslant \varepsilon\Big]\Big) = 0$$

ce qui prouve que pour tous réels $\lambda > 0$ et $\mu > 0$, $\lambda Z_n - \mu R_n$ est un estimateur convergent de $\lambda z(a,b) - \mu r(a,b)$.

L'énoncé incite alors explicitement à prendre $\lambda = \dfrac{2}{h-1} > 0$ et $\mu = \dfrac{2}{h(h-1)} > 0$, pour lesquels B_n défini au début de cette question, est donc un estimateur convergent de :

$$\begin{aligned}
\frac{2}{h-1}z(a,b) - \frac{2}{h(h-1)}r(a,b) &= \frac{2}{h-1}\ln\big(G_{a,b}(1)\big) - \frac{2}{h(h-1)}\ln\big(G_{a,b}(h)\big)\\
&= \frac{2}{h-1}\times\big(-a-\frac{b}{2}\big) - \frac{2}{h(h-1)}\times\big(-ah-\frac{b}{2}h^2\big)\\
&= \frac{-2a-b+2a+bh}{h-1}\\
&= b
\end{aligned}$$

On a bien démontré que B_n est un estimateur convergent du paramètre b.

★★★ FIN DU SUJET ★★★