Corrigés des sujets de révisions
Problèmes du TOP3

HEC/ESSEC I 2024 - nombre de triangles dans un graphe aléatoire, inégalité de Harris, inégalité de Janson, inégalité de Boole, espérance conditionnelle

Dans tout le sujet on considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, toutes les variables aléatoires qui interviennent dans la suite sont définies sur cet espace.

Soit n un entier supérieur ou égal à 3 et p un réel appartenant à $]0,1[$.

Pour générer des graphes non orientés de manière aléatoire, on se donne :

- $S = [\![0, n-1]\!]$, les sommets du graphe ;
- pour toute paire de sommets $\{u, v\}$ avec $u < v$, $T_{u,v}$ une variable aléatoire de Bernoulli de paramètre p.
  Les variables aléatoires $T_{u,v}$ pour $\{u, v\}$ décrivant les paires de sommets avec $u < v$, sont supposées indépendantes ;
- les arêtes d'un graphe G ainsi généré sont les paires $\{u, v\}$ telles que $T_{u,v} = 1$ si $u < v$ ou $T_{v,u} = 1$ si $v < u$.

Dans tout le problème, par convention, une somme portant sur un ensemble d'indices vide vaut 0, un produit vaut 1, une intersection vaut Ω, une réunion vaut $\varnothing$.

Partie 1 - Nombre aléatoire de triangles

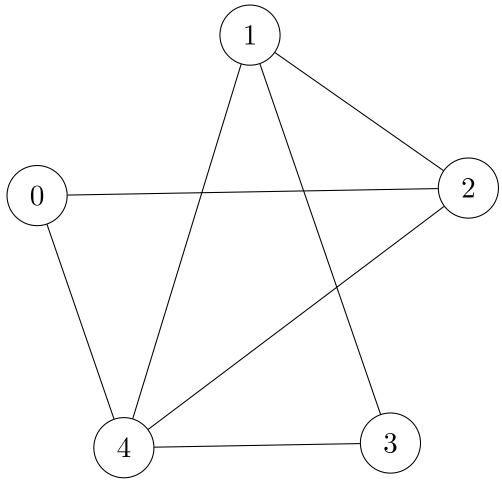
On note $\mathcal{T}$ l'ensemble des parties $\{u, v, w\}$ à trois éléments de l'ensemble des sommets, r le nombre de ses éléments et on pose

$$\mathcal{T} = \{t_1, \ldots, t_r\}$$

Étant donné $t = \{u, v, w\}$, un élément de $\mathcal{T}$, on dit que t est un triangle dans un graphe G généré aléatoirement si $\{u, v\}$, $\{v, w\}$ et $\{w, u\}$ sont des arêtes de G.

Pour tout $k \in [\![1, r]\!]$, on note Y_k la variable aléatoire de Bernoulli associée à l'événement « t_k est un triangle de G » et Z_n la variable aléatoire égale au nombre de triangles de G.

Par exemple si $n = 5$ et le graphe G est représenté ainsi,



alors $Z_5 = 3$.

1. Quelle est la valeur de r en fonction de n ?

Démonstration.
Le nombre r est, par définition, le cardinal de l'ensemble $\mathcal{T}$. On en déduit que r est le nombre de parties à 3 éléments d'un ensemble à n éléments (l'ensemble $S = [\![0, n-1]\!]$).

$$\boxed{\text{Ainsi : } r = \binom{n}{3}.}$$

□

2. a) Soit $k \in [\![1, r]\!]$. Posons $t_k = \{u, v, w\}$ avec $u < v < w$. Montrer que $Y_k = T_{u,v}T_{v,w}T_{u,w}$.

Démonstration.

- Tout d'abord, on sait que les variables Y_k, $T_{u,v}$, $T_{v,w}$ et $T_{u,w}$ sont toutes les quatre des variables aléatoires suivant une loi de Bernoulli.
D'où : $Y_k(\Omega) = T_{u,v}(\Omega) = T_{v,w}(\Omega) = T_{u,w}(\Omega) = \{0, 1\}$.
On en déduit : $(T_{u,v}T_{v,w}T_{u,w})(\Omega) = \{0, 1\} = Y_k(\Omega)$.
- Ensuite, en utilisant successivement les définitions de Y_k, d'un triangle et des variables aléatoires $T_{u,v}$, $T_{v,w}$ et $T_{u,w}$, on a, pour tout $\omega \in \Omega$:

$$\begin{aligned} Y_k(\omega) = 1 &\iff t_k \text{ est un triangle de } G \\ &\iff \{u, v\},\ \{v, w\} \text{ et } \{w, u\} \text{ sont des arêtes de } G \\ &\iff T_{u,v}(\omega) = 1 \text{ et } T_{v,w}(\omega) = 1 \text{ et } T_{u,w}(\omega) = 1 \qquad \textit{(car } u < v < w\textit{)} \\ &\iff (T_{u,v}T_{v,w}T_{u,w})(\omega) = 1 \end{aligned}$$

La dernière équivalence est justifiée (pour la réciproque) par le fait qu'un produit de réels est non nul si et seulement si tous les termes du produit sont non nuls.
Ainsi : $\forall \omega \in \Omega$, $Y_k(\omega) = (T_{u,v}T_{v,w}T_{u,w})(\omega)$ (car $Y_k(\Omega) = \{0, 1\} = (T_{u,v}T_{v,w}T_{u,w})(\Omega)$).

$$\boxed{\text{D'où : } Y_k = T_{u,v}T_{v,w}T_{u,w}.}$$

□

Commentaire

Rappelons qu'une égalité entre deux variables aléatoires est une égalité entre deux applications.
Ainsi, pour montrer que $X = Y$ (où X et Y sont deux variables aléatoires), on démontre :

$$\forall \omega \in \Omega, X(\omega) = Y(\omega)$$

b) En déduire que, pour tout $k \in [\![1, r]\!]$, Y_k suit la loi de Bernoulli de paramètre p^3.

Démonstration.
Soit $k \in [\![1, r]\!]$. D'après la question ***2.a)*** :

$$\begin{aligned} \mathbb{P}([Y_k = 1]) &= \mathbb{P}([T_{u,v}T_{v,w}T_{u,w} = 1]) \\ &= \mathbb{P}([T_{u,v} = 1] \cap [T_{v,w} = 1] \cap [T_{u,w} = 1]) \\ &= \mathbb{P}([T_{u,v} = 1])\mathbb{P}([T_{v,w} = 1])\mathbb{P}([T_{u,w} = 1]) \qquad \textit{(car } T_{u,v},\ T_{v,w} \textit{ et } T_{u,w} \textit{ sont indépendantes)} \\ &= p \times p \times p = p^3 \qquad \textit{(car } T_{u,v},\ T_{v,w} \textit{ et } T_{u,w} \textit{ suivent la loi } \mathcal{B}(p)\textit{)} \end{aligned}$$

Or, on sait déjà que Y_k suit une loi de Bernoulli. Ainsi, $\mathbb{P}([Y_k = 1])$ est le paramètre de cette loi.

$$\boxed{\text{On a bien : pour tout } k \in [\![1, r]\!],\ Y_k \hookrightarrow \mathcal{B}\left(p^3\right).}$$

□

c) Justifier que $Z_n = \sum\limits_{k=1}^{r} Y_k$. En déduire que $\mathbb{E}(Z_n) = \dbinom{n}{3} p^3$.

Démonstration.
Soit $\omega \in \Omega$. Rappelons :

$$Y_k(\omega) = \begin{cases} 1 & \text{si } t_k \text{ est un triangle} \\ 0 & \text{sinon} \end{cases}$$

La somme $\sum\limits_{k=1}^{r} Y_k(\omega)$ est alors une somme de 0 et de 1.
Plus précisément, si l'on note $J = \{k \in [\![1, r]\!] \mid t_k \text{ est un triangle}\} = \{k \in [\![1, r]\!] \mid Y_k(\omega) = 1\}$, alors :

$$\sum_{k=1}^{r} Y_k(\omega) = \sum_{k \in J} Y_k(\omega) = \sum_{k \in J} 1 = \text{Card}(J)$$

Or, $\text{Card}(J)$ est précisément le nombre de triangles de G, donc $\sum\limits_{k=1}^{r} Y_k(\omega) = Z_n(\omega)$.

$$\boxed{\text{Ceci étant vrai pour tout } \omega \in \Omega, \text{ on a bien : } Z_n = \sum_{k=1}^{r} Y_k.}$$

La variable aléatoire Z_n est finie ($Z_n(\Omega) = [\![0, r]\!]$) donc admet une espérance.

$$\begin{aligned}
\mathbb{E}(Z_n) &= \mathbb{E}\left(\sum_{k=1}^{r} Y_k\right) \\
&= \sum_{k=1}^{r} \mathbb{E}(Y_k) && \textit{(par linéarité de l'espérance)} \\
&= \sum_{k=1}^{r} p^3 && \textit{(d'après la question } \textbf{\textit{2.a)}}\textit{)} \\
&= r p^3 \\
&= \binom{n}{3} p^3 && \textit{(d'après la question } \textbf{\textit{1.}}\textit{)}
\end{aligned}$$

$$\boxed{\text{On a bien : } \mathbb{E}(Z_n) = \binom{n}{3} p^3.}$$

□

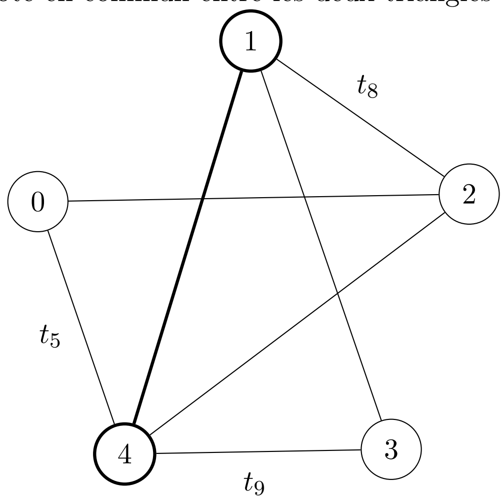
- ▶ On s'intéresse à la variance de Z_n.
  Si i et j appartiennent à $[\![1, r]\!]$ et sont différents, on note $i \equiv j$ lorsque t_i et t_j ont exactement deux éléments en commun et $i \not\equiv j$ dans le cas contraire.
  On note $\mathcal{E}$ l'ensemble des couples (i, j) tels que $i \equiv j$, et $\mathcal{F}$ l'ensemble des couples (i, j) tels que $i \neq j$ et $i \not\equiv j$.
  On désigne par a_n le nombre d'éléments de $\mathcal{E}$.

Commentaire

On peut regretter l'absence d'exemple pour aider à la compréhension de cette définition, donnée de manière abstraite. Tout d'abord, il faut s'empresser de penser que $i \equiv j$ lorsque t_i et t_j ont exactement deux **sommets** en commun, ce qui permet de faire appel à l'intuition géométrique habituelle sur la notion de triangle. En effet, si $t = \{u, v, w\}$ est un triangle de G, alors un élément de t (qui ne peut être que u, v ou w) est un sommet du graphe G et également un sommet du triangle t. De plus, considérant deux triangles t_i et t_j, avoir exactement deux sommets en commun, c'est avoir exactement un **côté** en commun (en effet, si deux côtés sont égaux alors le troisième aussi et donc les trois sommets sont égaux).
Reprenons le graphe à 5 sommets donné en exemple introductif et choisissons de numéroter les éléments de $\mathcal{T}$ en utilisant l'ordre lexicographique :

$$\begin{array}{lllll} t_1 = \{0,1,2\}, & t_2 = \{0,1,3\}, & t_3 = \{0,1,4\}, & t_4 = \{0,2,3\}, & t_5 = \{0,2,4\}, \\ t_6 = \{0,3,4\}, & t_7 = \{1,2,3\}, & t_8 = \{1,2,4\}, & t_9 = \{1,3,4\}, & t_{10} = \{2,3,4\}. \end{array}$$

On représente à nouveau le graphe G, en faisant apparaître les trois triangles et en épaississant le trait pour les sommets et le côté en commun entre les deux triangles t_8 et t_9 :



On a donc $8 \equiv 9$ sur cet exemple. C'est le seul couple (i, j) dans $\mathcal{E}$ tel que t_i et t_j soient des triangles, mais il y a d'autres couples (i, j) dans $\mathcal{E}$. Par exemple : $1 \equiv 2$, $4 \equiv 6$,...

3. a) Montrer que :

$$\mathbb{V}(Z_n) = \operatorname{Cov}\left(\sum_{i=1}^{r} Y_i, \sum_{j=1}^{r} Y_j\right) = \sum_{(i,j)\in[\![1,r]\!]^2} \operatorname{Cov}(Y_i, Y_j)$$

Démonstration.
Tout d'abord, la variable aléatoire Z_n est finie donc admet une variance.
D'après la formule de Koenig-Huygens :

$$\mathbb{V}(Z_n) = \operatorname{Cov}(Z_n, Z_n)$$

et d'après la question ***2.c)*** : $Z_n = \sum\limits_{i=1}^{r} Y_i$. D'où :

$$\mathbb{V}(Z_n) = \operatorname{Cov}\left(\sum_{i=1}^{r} Y_i, \sum_{j=1}^{r} Y_j\right)$$

Ensuite :

$$\begin{aligned}
\mathrm{Cov}\left(\sum_{i=1}^{r} Y_i, \sum_{j=1}^{r} Y_j\right) &= \sum_{i=1}^{r} \mathrm{Cov}\left(Y_i, \sum_{j=1}^{r} Y_j\right) && \textit{(par linéarité à gauche de } \mathrm{Cov}\textit{)} \\
&= \sum_{i=1}^{r} \sum_{j=1}^{r} \mathrm{Cov}\left(Y_i, Y_j\right) && \textit{(par linéarité à droite de } \mathrm{Cov}\textit{)} \\
&= \sum_{(i,j)\in[\![1,r]\!]^2} \mathrm{Cov}(Y_i, Y_j)
\end{aligned}$$

$$\boxed{\text{On a bien : } \mathbb{V}(Z_n) = \mathrm{Cov}\left(\sum_{i=1}^{r} Y_i, \sum_{j=1}^{r} Y_j\right) = \sum_{(i,j)\in[\![1,r]\!]^2} \mathrm{Cov}(Y_i, Y_j).}$$

□

b) Montrer que si $(i,j) \in \mathcal{F}$, Y_i et Y_j sont indépendantes. En déduire que :

$$\mathbb{V}(Z_n) = \sum_{i=1}^{r} \mathbb{V}(Y_i) + \sum_{(i,j)\in\mathcal{E}} \mathrm{Cov}(Y_i, Y_j)$$

Démonstration.

- Soit $(i,j) \in \mathcal{F}$. Alors $i \neq j$ et $i \not\equiv j$.
  Notons $t_i = \{u_i, v_i, w_i\}$ avec $u_i < v_i < w_i$ et $t_j = \{u_j, v_j, w_j\}$ avec $u_j < v_j < w_j$.
  D'après la question ***2.a)*** :

  $$Y_i = T_{u_i,v_i} T_{v_i,w_i} T_{u_i,w_i} \qquad \text{et} \qquad Y_j = T_{u_j,v_j} T_{v_j,w_j} T_{u_j,w_j}$$

  × Puisque $i \neq j$, on peut affirmer que $t_i \neq t_j$, ce qui veut dire que t_i et t_j ne partagent pas trois sommets en commun.

  × Puisque $i \not\equiv j$, on peut affirmer que t_i et t_j ne partagent pas exactement deux sommets en commun.

  On en déduit que t_i et t_j ont soit 0, soit 1 sommet en commun. Ainsi, les variables aléatoires $T_{u_i,v_i}, T_{v_i,w_i}, T_{u_i,w_i}, T_{u_j,v_j}, T_{v_j,w_j}$ et T_{u_j,w_j} sont distinctes, ce qui permet de conclure qu'elles sont mutuellement indépendantes.

  $$\boxed{\text{D'après le lemme des coalitions : } Y_i \text{ et } Y_j \text{ sont indépendantes.}}$$

- Reprenons le calcul de la question ***3.a)*** :

  $$\begin{aligned}
  \mathbb{V}(Z_n) &= \sum_{(i,j)\in[\![1,r]\!]^2} \mathrm{Cov}(Y_i, Y_j) \\
  &= \sum_{\substack{(i,j)\in[\![1,r]\!]^2 \\ i=j}} \mathrm{Cov}(Y_i, Y_j) + \sum_{\substack{(i,j)\in[\![1,r]\!]^2 \\ i\neq j}} \mathrm{Cov}(Y_i, Y_j)
  \end{aligned}$$

  × D'une part :

  $$\sum_{\substack{(i,j)\in[\![1,r]\!]^2 \\ i=j}} \mathrm{Cov}(Y_i, Y_j) = \sum_{i=1}^{r} \mathrm{Cov}(Y_i, Y_i) = \sum_{i=1}^{r} \mathbb{V}(Y_i)$$

  × D'autre part, par définition de $\mathcal{E}$ et $\mathcal{F}$, on a :

  $$\mathcal{E} \cup \mathcal{F} = \{(i,j) \in [\![1,r]\!]^2 \mid i \neq j\}$$

  et il s'agit d'une union disjointe.

Ainsi :

$$\sum_{\substack{(i,j)\in[\![1,r]\!]^2\\ i\neq j}} \text{Cov}(Y_i,Y_j) = \sum_{(i,j)\in\mathcal{E}} \text{Cov}(Y_i,Y_j) + \sum_{(i,j)\in\mathcal{F}} \text{Cov}(Y_i,Y_j)$$

De plus, pour tout $(i,j)\in\mathcal{F}$, Y_i et Y_j sont indépendantes et donc $\text{Cov}(Y_i,Y_j)=0$. Il suit :

$$\sum_{(i,j)\in\mathcal{F}} \text{Cov}(Y_i,Y_j) = 0$$

D'où :

$$\sum_{\substack{(i,j)\in[\![1,r]\!]^2\\ i\neq j}} \text{Cov}(Y_i,Y_j) = \sum_{(i,j)\in\mathcal{E}} \text{Cov}(Y_i,Y_j)$$

$$\boxed{\text{On a bien : } \mathbb{V}(Z_n) = \sum_{i=1}^{r} \mathbb{V}(Y_i) + \sum_{(i,j)\in\mathcal{E}} \text{Cov}(Y_i,Y_j).}$$

□

c) En conclure que :

$$\mathbb{V}(Z_n) = rp^3(1-p^3) + \left(\sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j)\right) - a_np^6$$

Démonstration.

- Tout d'abord, d'après la question ***2.b)*** : pour tout $i\in[\![1,r]\!]$, $Y_i \hookrightarrow \mathcal{B}\left(p^3\right)$.
  Donc : pour tout $i\in[\![1,r]\!]$, $\mathbb{V}(Y_i)=p^3(1-p^3)$.
  On en déduit :
  $$\sum_{i=1}^{r} \mathbb{V}(Y_i) = \sum_{i=1}^{r} p^3(1-p^3) = rp^3(1-p^3)$$
- Ensuite, d'après la formule de Koenig-Huygens, pour tout $(i,j)\in\mathcal{E}$:
  $$\begin{aligned}\text{Cov}(Y_i,Y_j) &= \mathbb{E}(Y_iY_j) - \mathbb{E}(Y_i)\mathbb{E}(Y_j)\\ &= \mathbb{E}(Y_iY_j) - p^3p^3\\ &= \mathbb{E}(Y_iY_j) - p^6\end{aligned}$$

  Ainsi :
  $$\begin{aligned}\sum_{(i,j)\in\mathcal{E}} \text{Cov}(Y_i,Y_j) &= \sum_{(i,j)\in\mathcal{E}} \left(\mathbb{E}(Y_iY_j)-p^6\right)\\ &= \sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j) - \sum_{(i,j)\in\mathcal{E}} p^6 && \textit{(par linéarité de la somme)}\\ &= \left(\sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j)\right) - a_np^6 && \textit{(car } a_n = \text{Card}(\mathcal{E})\textit{)}\end{aligned}$$

$$\boxed{\text{D'après la question } \textbf{\textit{3.b)}} : \mathbb{V}(Z_n) = rp^3(1-p^3) + \left(\sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j)\right) - a_np^6.}$$

□

▶ On note $\Delta_n = \sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j)$

4. Montrer que si $i \equiv j$, $\mathbb{E}(Y_iY_j) = p^5$ et en déduire que $\Delta_n = a_np^5$.
En conclure que : $\mathbb{V}(Z_n) = \binom{n}{3}(p^3 - p^6) + a_n(p^5 - p^6)$.

Démonstration.

- Soit $(i, j) \in [\![1, r]\!]^2$ tel que $i \equiv j$. On a en particulier $i \neq j$.
Notons $t_i = \{u_i, v_i, w_i\}$ avec $u_i < v_i < w_i$ et $t_j = \{u_j, v_j, w_j\}$ avec $u_j < v_j < w_j$.
D'après la question ***2.a)*** :

$$Y_i = T_{u_i,v_i}T_{v_i,w_i}T_{u_i,w_i} \qquad \text{et} \qquad Y_j = T_{u_j,v_j}T_{v_j,w_j}T_{u_j,w_j}$$

Par définition de la relation $\equiv$, les triangles t_i et t_j ont exactement deux sommets en commun. La paire de sommets en commun ne peut être que l'une des trois suivantes :

$$\{u_i, v_i\} \qquad \texttt{OU} \qquad \{v_i, w_i\} \qquad \texttt{OU} \qquad \{u_i, w_i\}$$

Supposons que la paire de sommets en commun soit $\{u_i, v_i\}$ pour fixer les idées (les autres cas sont analogues). On a alors :

$$\begin{aligned} Y_iY_j &= T_{u_i,v_i}T_{v_i,w_i}T_{u_i,w_i}T_{u_j,v_j}T_{v_j,w_j}T_{u_j,w_j} \\ &= T_{v_i,w_i}T_{u_i,w_i}T_{u_j,v_j}T_{v_j,w_j}T_{u_j,w_j} && \textit{(car } T_{u_i,v_i} \textit{ apparaît deux fois dans ce produit et } T_{u_i,v_i}^2 = T_{u_i,v_i} \textit{ puisque } T_{u_i,v_i}(\Omega) = \{0,1\}\textit{)} \end{aligned}$$

et on remarque que les variables restantes sont deux à deux distinctes (sinon les triangles t_i et t_j auraient trois sommets en commun, mais cela contredirait le fait que $i \neq j$) et donc sont mutuellement indépendantes. Puisque elles admettent toutes une espérance égale à p, il vient :

$$\begin{aligned} \mathbb{E}(Y_iY_j) &= \mathbb{E}(T_{v_i,w_i}T_{u_i,w_i}T_{u_j,v_j}T_{v_j,w_j}T_{u_j,w_j}) \\ &= \mathbb{E}(T_{v_i,w_i})\mathbb{E}(T_{u_i,w_i})\mathbb{E}(T_{u_j,v_j})\mathbb{E}(T_{v_j,w_j})\mathbb{E}(T_{u_j,w_j}) && \textit{(par indépendance des variables aléatoires en présence)} \\ &= p^5 \end{aligned}$$

- D'après ce qui précède :

$$\sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j) = \sum_{(i,j)\in\mathcal{E}} p^5 = a_np^5 \qquad \textit{(car } a_n = \mathrm{Card}(\mathcal{E})\textit{)}$$

$$\boxed{\text{On a bien : } \Delta_n = a_np^5.}$$

- On peut alors finir le calcul :

$$\begin{aligned} \mathbb{V}(Z_n) &= rp^3(1-p^3) + \left(\sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j)\right) - a_np^6 && \textit{(d'après la question } \textbf{3.c)}\textit{)} \\ &= rp^3(1-p^3) + a_np^5 - a_np^6 && \textit{(d'après ce qui précède)} \\ &= \binom{n}{3}(p^3 - p^6) + a_n(p^5 - p^6) && \textit{(d'après la question } \textbf{1.}\textit{)} \end{aligned}$$

$$\boxed{\text{On obtient bien : } \mathbb{V}(Z_n) = \binom{n}{3}(p^3 - p^6) + a_n(p^5 - p^6).}$$

□

Commentaire

Nous avons utilisé au cours de cette preuve le fait suivant : si X suit une loi de Bernoulli, alors $X^2 = X$.
Ce résultat n'est pas explicitement au programme, mais il est bon de le connaître. Il est probable que son utilisation au cours de cette preuve sans donner de justification détaillée ne soit pas sanctionnée puisque ce n'est pas le cœur de la preuve. Détaillons toutefois l'argument.
Soit $\omega \in \Omega$. Deux cas se présentent.

× Premier cas : $X(\omega) = 1$.
Alors $X^2(\omega) = 1^2 = 1 = X(\omega)$.

× Deuxième cas : $X(\omega) = 0$.
Alors $X^2(\omega) = 0^2 = 0 = X(\omega)$.

On a donc montré : $\forall \omega \in \Omega, X^2(\omega) = X(\omega)$. Ceci prouve que $X^2 = X$.
Une fois résumée, la preuve tient en une phrase : les valeurs possibles pour X sont toutes des points fixes de la fonction $x \mapsto x^2$.

5. Calcul de a_n.

a) Déterminer le nombre de triplets $(\{u, v\}, w, y)$ où u, v, w, y sont quatre éléments distincts de l'ensemble $[\![0, n-1]\!]$.

Démonstration.
Un tel triplet est entièrement déterminé par :

× le choix de la paire $\{u, v\}$: $\binom{n}{2} = \dfrac{n(n-1)}{2}$ possibilités,

× le choix de l'élément w (distinct de u et v) : $\binom{n-2}{1} = n-2$ possibilités,

× le choix de l'élément y (distinct de u, v et w) : $\binom{n-3}{1} = n-3$ possibilités.

$$\boxed{\text{Il y a donc } \dfrac{n(n-1)(n-2)(n-3)}{2} \text{ tels triplets.}}$$

□

b) En déduire que $a_n = \dfrac{n(n-1)(n-2)(n-3)}{2}$.

Démonstration.
Soit $(i, j) \in \mathcal{E}$. Le couple (i, j) est entièrement déterminé par :

× le choix de la paire $\{u, v\}$ de sommets communs aux triangles t_i et t_j,

× le choix de l'élément w (distinct de u et v) qui sera le troisième et dernier sommet de t_i,

× le choix de l'élément y (distinct de u, v et w) qui sera le troisième et dernier sommet de t_j.

D'après la question ***5.a)***, il y a $\dfrac{n(n-1)(n-2)(n-3)}{2}$ tels triplets $(\{u, v\}, w, y)$.

$$\boxed{\text{On en déduit que } a_n = \dfrac{n(n-1)(n-2)(n-3)}{2}.}$$

□

Commentaire

Notons $E = \{(\{u,v\},w,y) \mid u,v,w,y$ sont quatre éléments distincts de l'ensemble $[\![0,n-1]\!]\}$.
D'après la question ***5.a)*** : $\text{Card}(E) = \dfrac{n(n-1)(n-2)(n-3)}{2}$.
On peut rendre la preuve précédente plus rigoureuse en construisant une bijection de $\mathcal{E}$ vers E.

- Soit $(i,j) \in \mathcal{E}$.
  Notons u et v les deux sommets (distincts) qui sont communs aux triangles t_i et t_j puis notons w (resp. y) le troisième sommet de t_i (resp. de t_j). Remarquons alors que $t_i = \{u,v,w\}$ et $t_j = \{u,v,y\}$.
  Les objets $\{u,v\}$, w et y sont entièrement déterminés par le choix du couple $(i,j) \in \mathcal{E}$ (puisque nous n'avons pas fixé d'ordre entre u et v) et on a bien $(\{u,v\},w,y) \in E$.
  On en déduit que l'on peut définir, avec ces notations, l'application :
  $$\Phi : \left\{ \begin{array}{ccc} \mathcal{E} & \to & E \\ (i,j) & \mapsto & (\{u,v\},w,y) \end{array} \right.$$
- Réciproquement, soit $(\{u,v\},w,y) \in E$.
  Notons i (resp. j) l'unique entier de $[\![1,r]\!]$ vérifiant $t_i = \{u,v,w\}$ (resp. $t_j = \{u,v,y\}$).
  Le couple (i,j) est entièrement déterminé par le choix du triplet $(\{u,v\},w,y) \in E$ et on a bien $(i,j) \in \mathcal{E}$ (les deux triangles ont exactement deux sommets en commun : u et v).
  On en déduit que l'on peut définir, avec ces notations, l'application :
  $$\Psi : \left\{ \begin{array}{ccc} E & \to & \mathcal{E} \\ (\{u,v\},w,y) & \mapsto & (i,j) \end{array} \right.$$
- On vérifie alors que $\Psi \circ \Phi = \text{id}_{\mathcal{E}}$ et $\Phi \circ \Psi = \text{id}_E$ (ce qui est évident par construction). Cela prouve que $\mathcal{E}$ et E sont en bijection. On en déduit qu'ils possèdent le même cardinal et donc : $a_n = \dfrac{n(n-1)(n-2)(n-3)}{2}$.

Partie 2 - Étude informatique

On se donne un graphe G généré par le procédé décrit dans le préambule.

On définit la fonction `supprimeDer(L)` qui si `L` est la liste des listes d'adjacence du graphe G dont les sommets sont $0, 1, \dots, n-1$, modifie `L` afin qu'elle devienne la liste des listes d'adjacence du graphe G', dont les sommets sont $0, 1, \dots, n-2$, obtenu en supprimant dans G le sommet $n-1$ et les arêtes contenant ce sommet.

```python
def supprimeDer(L):
    s = len(L) - 1
    L.pop() # supprime le dernier élément de la liste L
    for a in L:
        if s in a:
            a.remove(s) # supprime s dans la liste a
```

6. Compléter la fonction suivante pour qu'elle retourne le nombre de triangles dont un des sommets est le sommet `s` dans le graphe G dont la liste des listes d'adjacence est `L` :

```
def triangle2s(s, L):
    cpt = 0
    adj = L[s]
    for i in range(len(adj)):
        for j in range(..., len(adj)) :
            if ... in L[...]:
                cpt += 1
    return cpt
```

Démonstration.
Considérons $x < y$ deux voisins (distincts) de s.
Le programme qui suit est basé sur la remarque suivante : le triplet $\{s, x, y\}$ est un triangle si et seulement si x et y sont voisins.

```
1 def triangle2s(s, L):
2     cpt = 0
3     adj = L[s] # liste des voisins de s
4     # on teste une et une seule fois tous les couples de voisins de s
5     for i in range(len(adj)):
6         for j in range(i+1, len(adj)) :
7             if adj[j] in L[adj[i]]: # si adj[i] et adj[j] sont voisins
8                 cpt += 1
9     return cpt
```

Les lignes 5 et 6 permettent de considérer tous les couples (x, y) de voisins de s vérifiant $x < y$, avec `x = adj[i]` et `y = adj[j]`. □

7. Écrire une fonction `nbTriangles(L)`, utilisant les deux fonctions précédentes, qui retourne le nombre de triangles du graphe G dont la liste des listes d'adjacence est représentée par L.

Démonstration.
L'idée est de compter les triangles dont l'un des sommets est le dernier sommet du graphe G (à l'aide de la fonction `triangle2s(s, L)`) puis de supprimer ce dernier sommet (à l'aide de la fonction `supprimeDer(L)`) puis de recommencer jusqu'à-ce que le graphe ne comporte plus que trois sommets (un graphe à deux sommets ne peut pas contenir de triangles). A la fin, on somme tous les nombres obtenus. De cette manière, tous les triangles sont pris en compte et on ne compte jamais deux fois le même triangle.

```
def nbTriangles(L):
    cpt = 0 # compteur du nombre total de triangles dans le graphe G
    s = len(L) - 1 # s est le dernier sommet du graphe G
    while s >= 2: # cas limite : graphe à trois sommets (0, 1 et 2)
        cpt += triangle2s(s, L) # ajout du nombre de triangles ayant s pour sommet
        supprimeDer(L) # obtention du nouveau graphe en enlevant s
        s = s-1 # mise à jour de s pour que ce soit toujours le dernier sommet
    return cpt
```

□

8. On suppose que la fonction `graphe(n,p)` génère un graphe aléatoire suivant les hypothèses décrites dans le préambule.
Expliquer ce que retourne la fonction suivante :

```
1 def fonctionMystere(n):
2     cpt = 0
3     for i in range(1000):
4         L = graphe(n, 1/n)
5         if nbTriangles(L) == 0:
6             cpt += 1
7     return cpt / 1000
```

Démonstration. Détaillons ce que fait cette fonction.

- A la ligne 2, une variable `cpt` est initialisée à 0. Cette variable servira de compteur.
- Les lignes 3 et 4 consistent à simuler 1000 graphes aléatoires de paramètres n et $p = \frac{1}{n}$.
- Pour chaque graphe aléatoire généré, on teste si ce graphe ne contient aucun triangle (ligne 5). Si c'est bien le cas, on incrémente la variable `cpt` (ligne 6).

Ainsi, en fin de boucle `for`, la variable `cpt` contient le nombre de graphes aléatoires ne contenant aucun triangle, c'est-à-dire le nombre de fois où l'événement $[Z_n = 0]$ a été réalisé.
On en déduit que la fonction `fonctionMystere(n)` renvoie la fréquence empirique de réalisation de l'événement $[Z_n = 0]$ (dans le cas spécifique où $p = \frac{1}{n}$). Cette fréquence empirique n'est rien d'autre qu'une réalisation de la moyenne empirique $\overline{W}_{1000} = \frac{1}{1000} \sum_{i=1}^{1000} W_i$ où $(W_1, \dots, W_{1000})$ est un 1000-échantillon d'une variable aléatoire W qui prend la valeur 1 si l'événement $[Z_n = 0]$ est réalisé, et la valeur 0 sinon.

D'après la loi faible des grands nombres, la fonction `fonctionMystere(n)` renvoie donc une approximation de $\mathbb{P}([Z_n = 0])$. □

Partie 3 - Inégalité de Harris

k désigne un entier naturel non nul.

- Soit f une fonction définie sur une partie $\mathcal{D}$ de $\mathbb{R}^k$ à valeurs dans $\mathbb{R}$.
  Si $k \geqslant 2$, on dit que f est k-croissante sur $\mathcal{D}$ si, pour tout $(x_1, \dots, x_k)$ élément de $\mathcal{D}$ et $i \in [\![1, k]\!]$, $t \mapsto f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_k)$ est croissante sur son ensemble de définition.
  Si $k = 1$, une fonction 1-croissante sur $\mathcal{D}$ est simplement une fonction croissante sur $\mathcal{D}$.
- On définit de même la notion de fonction k-décroissante.
- On considère $X_1, \dots, X_k$ des variables aléatoires finies.
  On admet le résultat suivant (théorème de transfert d'ordre k) :
  Si f est une fonction définie sur $X_1(\Omega) \times \cdots \times X_k(\Omega)$ et $Y_k = f(X_1, \dots, X_k)$ alors

$$\mathbb{E}(Y_k) = \sum_{(x_1,\dots,x_k) \in X_1(\Omega) \times \cdots \times X_k(\Omega)} f(x_1, \dots, x_k) \mathbb{P}\left([X_1 = x_1] \cap \cdots \cap [X_k = x_k]\right)$$

On note (H_k) la propriété suivante :

Si $X_1, \dots, X_k$ sont des variables aléatoires finies indépendantes, f et g deux fonctions définies sur $X_1(\Omega) \times \cdots \times X_k(\Omega)$ et k-croissantes sur cet ensemble, et si l'on pose $Y_k = f(X_1, \dots, X_k)$ et $Z_k = g(X_1, \dots, X_k)$, alors :

$$\mathbb{E}(Y_k Z_k) \geqslant \mathbb{E}(Y_k)\mathbb{E}(Z_k) \quad \text{(inégalité de Harris)}$$

Commentaire

Ici aussi, on peut regretter l'absence d'un exemple pour aider les candidats à intégrer mentalement la définition proposée.
Dans le cas particulier où $k = 1$, où X_1 est une variable aléatoire finie et où $f = g = \mathrm{id}_{\mathbb{R}}$ (qui est bien croissante sur $\mathbb{R}$), on retrouve $\mathbb{E}(X_1^2) \geqslant \mathbb{E}(X_1)^2$, c'est-à-dire $\mathbb{V}(X_1) \geqslant 0$.

Commentaire

Cette partie fait un choix de notations un peu malheureux. La notation Y_k est utilisée de manière abstraite pour désigner une transformée des $X_1, \ldots, X_k$ alors que Y_k a été définie en début de partie 1 comme la variable aléatoire de Bernoulli associée à l'événement « t_k est un triangle de G ». Il fallait comprendre que la partie 3 est totalement indépendante du début du sujet.

9. Dans cette question $k = 1$, on pose $X = X_1$ une variable aléatoire finie, f et g sont deux fonctions croissantes sur $X(\Omega)$.

a) Montrer que pour tout $(x, y) \in (X(\Omega))^2$, $(f(x) - f(y))(g(x) - g(y)) \geqslant 0$.

Démonstration.
Soit $(x, y) \in (X(\Omega))^2$. Deux cas se présentent.

× Premier cas : $x \leqslant y$.
Alors $f(x) \leqslant f(y)$ et $g(x) \leqslant g(y)$ par croissance de f et g sur $X(\Omega)$.
On en déduit : $f(x) - f(y) \leqslant 0$ et $g(x) - g(y) \leqslant 0$.
D'où : $(f(x) - f(y))(g(x) - g(y)) \geqslant 0$.

× Deuxième cas : $x > y$.
Alors $f(x) \geqslant f(y)$ et $g(x) \geqslant g(y)$ par croissance de f et g sur $X(\Omega)$.
On en déduit : $f(x) - f(y) \geqslant 0$ et $g(x) - g(y) \geqslant 0$.
D'où : $(f(x) - f(y))(g(x) - g(y)) \geqslant 0$.

$$\boxed{\text{Dans tous les cas : } (f(x) - f(y))(g(x) - g(y)) \geqslant 0.}$$

□

Commentaire

Il ne faut jamais se décourager au cours de l'épreuve car une question facile et indépendante du reste de l'énoncé peut toujours arriver après plusieurs questions (trop) difficiles. C'est le cas de cette question ***9.a)*** qui est un résultat d'analyse reposant uniquement sur une définition du cours de première année sur les fonctions. Une très bonne connaissance du cours est une condition *sine qua non* de réussite aux concours, quelque soit le niveau de l'épreuve.

b) Montrer que pour tout $y \in X(\Omega)$,

$$\mathbb{E}(f(X)g(X)) + f(y)g(y) \geqslant g(y)\mathbb{E}(f(X)) + f(y)\mathbb{E}(g(X))$$

Démonstration.
D'après la question ***9.a)*** : pour tout $(x, y) \in (X(\Omega))^2$, $(f(x) - f(y))(g(x) - g(y)) \geqslant 0$.
Soit $y \in X(\Omega)$ et soit $\omega \in \Omega$. En appliquant l'inégalité précédente à $x = X(\omega) \in X(\Omega)$, on obtient :

$$(f(X(\omega)) - f(y))(g(X(\omega)) - g(y)) \geqslant 0$$

Ainsi :
$$(f(X) - f(y))(g(X) - g(y)) \geqslant 0$$
En développant cette expression, on obtient :
$$f(X)g(X) + f(y)g(y) \geqslant g(y)f(X) + f(y)g(X)$$
Toutes les variables aléatoires en présence sont finies et donc admettent une espérance.
Par croissance de l'espérance :
$$\mathbb{E}(f(X)g(X) + f(y)g(y)) \geqslant \mathbb{E}(g(y)f(X) + f(y)g(X))$$
Par linéarité de l'espérance :
$$\boxed{\mathbb{E}(f(X)g(X)) + f(y)g(y) \geqslant g(y)\mathbb{E}(f(X)) + f(y)\mathbb{E}(g(X))}$$
□

c) En déduire que (H_1) est vraie.

Démonstration.
Soit $\omega \in \Omega$. D'après la question ***9.b)*** appliquée à $y = X(\omega) \in X(\Omega)$:
$$\mathbb{E}(f(X)g(X)) + f(X(\omega))g(X(\omega)) \geqslant g(X(\omega))\mathbb{E}(f(X)) + f(X(\omega))\mathbb{E}(g(X))$$
On en déduit :
$$\mathbb{E}(f(X)g(X)) + f(X)g(X) \geqslant g(X)\mathbb{E}(f(X)) + f(X)\mathbb{E}(g(X))$$
Par croissance de l'espérance :
$$\mathbb{E}\big(\mathbb{E}(f(X)g(X)) + f(X)g(X)\big) \geqslant \mathbb{E}\big(g(X)\mathbb{E}(f(X)) + f(X)\mathbb{E}(g(X))\big)$$
Par linéarité de l'espérance :
$$\mathbb{E}(f(X)g(X)) + \mathbb{E}(f(X)g(X)) \geqslant \mathbb{E}(f(X))\mathbb{E}(g(X)) + \mathbb{E}(g(X))\mathbb{E}(f(X))$$
et donc :
$$2\mathbb{E}(f(X)g(X)) \geqslant 2\mathbb{E}(f(X))\mathbb{E}(g(X))$$
d'où :
$$\mathbb{E}(f(X)g(X)) \geqslant \mathbb{E}(f(X))\mathbb{E}(g(X))$$
$$\boxed{\text{La propriété } (H_1) \text{ est vraie.}}$$
□

10. On suppose que (H_k) est vraie pour un certain k et on considère $X_1, \dots, X_{k+1}$ des variables aléatoires finies indépendantes, f et g deux fonctions définies sur $X_1(\Omega) \times \dots \times X_{k+1}(\Omega)$ et $(k+1)$-croissantes.
On pose $Y_{k+1} = f(X_1, \dots, X_{k+1})$ et $Z_{k+1} = g(X_1, \dots, X_{k+1})$.

a) A l'aide des théorèmes de transfert d'ordre $k+1$ et k, montrer que :
$$\mathbb{E}(Y_{k+1}Z_{k+1}) = \sum_{x \in X_{k+1}(\Omega)} \mathbb{E}(f(X_1, \dots, X_k, x)g(X_1, \dots, X_k, x))\mathbb{P}([X_{k+1} = x])$$

Démonstration.
D'après le théorème de transfert d'ordre $k+1$ appliqué à la fonction
$h_{k+1} : (x_1, \dots, x_{k+1}) \mapsto f(x_1, \dots, x_{k+1}) g(x_1, \dots, x_{k+1})$, on a :

$$\mathbb{E}(Y_{k+1} Z_{k+1})$$

$$= \mathbb{E}(h_{k+1}(X_1, \dots, X_{k+1}))$$

$$= \sum_{\substack{(x_1,\dots,x_{k+1}) \in \\ X_1(\Omega) \times \dots \times X_{k+1}(\Omega)}} f(x_1, \dots, x_{k+1}) g(x_1, \dots, x_{k+1}) \mathbb{P}\left([X_1 = x_1] \cap \dots \cap [X_{k+1} = x_{k+1}]\right)$$

$$= \sum_{x \in X_{k+1}(\Omega)} \sum_{\substack{(x_1,\dots,x_k) \in \\ X_1(\Omega) \times \dots \times X_k(\Omega)}} f(x_1, \dots, x_k, x) g(x_1, \dots, x_k, x) \mathbb{P}\left([X_1 = x_1] \cap \dots \cap [X_k = x_k] \cap [X_{k+1} = x]\right)$$

(car x_k est une variable muette)

$$= \sum_{x \in X_{k+1}(\Omega)} \left(\sum_{\substack{(x_1,\dots,x_k) \in \\ X_1(\Omega) \times \dots \times X_k(\Omega)}} f(x_1, \dots, x_k, x) g(x_1, \dots, x_k, x) \mathbb{P}\left(\bigcap_{i=1}^{k} [X_i = x_i] \right) \right) \mathbb{P}([X_{k+1} = x])$$

où l'on a utilisé l'indépendance mutuelle de $X_1, \dots, X_{k+1}$ à la dernière ligne.
Soit $x \in X_{k+1}(\Omega)$. D'après le théorème de transfert d'ordre k appliqué à la fonction
$h_{k,x} : (x_1, \dots, x_k) \mapsto f(x_1, \dots, x_k, x) g(x_1, \dots, x_k, x)$, on a :

$$\mathbb{E}(f(X_1, \dots, X_k, x) g(X_1, \dots, X_k, x)) = \sum_{\substack{(x_1,\dots,x_k) \in \\ X_1(\Omega) \times \dots \times X_k(\Omega)}} f(x_1, \dots, x_k, x) g(x_1, \dots, x_k, x) \mathbb{P}\left([X_1 = x_1] \cap \dots \cap [X_k = x_k]\right)$$

$$\boxed{\text{D'où : } \mathbb{E}(Y_{k+1} Z_{k+1}) = \sum_{x \in X_{k+1}(\Omega)} \mathbb{E}(f(X_1, \dots, X_k, x) g(X_1, \dots, X_k, x)) \mathbb{P}([X_{k+1} = x]).}$$ □

b) Justifier que pour tout $x \in X_{k+1}(\Omega)$:

$$\mathbb{E}(f(X_1, \dots, X_k, x) g(X_1, \dots, X_k, x)) \geqslant \mathbb{E}(f(X_1, \dots, X_k, x)) \mathbb{E}(g(X_1, \dots, X_k, x))$$

Démonstration.
Soit $x \in X_{k+1}(\Omega)$.
Notons $\tilde{f} : (x_1, \dots, x_k) \mapsto f(x_1, \dots, x_k, x)$ de sorte que $f(X_1, \dots, X_k, x) = \tilde{f}(X_1, \dots, X_k)$.
Notons $\tilde{g} : (x_1, \dots, x_k) \mapsto g(x_1, \dots, x_k, x)$ de sorte que $g(X_1, \dots, X_k, x) = \tilde{g}(X_1, \dots, X_k)$.

× Les variables $X_1, \dots, X_k$ sont indépendantes.

× Les fonctions $\tilde{f}$ et $\tilde{g}$ sont définies et k-croissantes sur $X_1(\Omega) \times \dots \times X_k(\Omega)$ car les fonctions f et g sont définies et $(k+1)$-croissantes sur $X_1(\Omega) \times \dots \times X_{k+1}(\Omega)$.

× On pose $\tilde{Y}_k = \tilde{f}(X_1, \dots, X_k)$ et $\tilde{Z}_k = \tilde{g}(X_1, \dots, X_k)$.

On a supposé que (H_k) était vraie, donc on a :

$$\mathbb{E}(\tilde{Y}_k \tilde{Z}_k) \geqslant \mathbb{E}(\tilde{Y}_k) \mathbb{E}(\tilde{Z}_k)$$

Autrement dit :

$$\boxed{\mathbb{E}(f(X_1, \dots, X_k, x) g(X_1, \dots, X_k, x)) \geqslant \mathbb{E}(f(X_1, \dots, X_k, x)) \mathbb{E}(g(X_1, \dots, X_k, x)).}$$ □

c) On pose pour tout $x \in X_{k+1}(\Omega)$, $u(x) = \mathbb{E}(f(X_1, \dots, X_k, x))$ et $v(x) = \mathbb{E}(g(X_1, \dots, X_k, x))$.
Montrer que u et v sont croissantes sur $X_{k+1}(\Omega)$ et $\mathbb{E}(Y_{k+1} Z_{k+1}) \geqslant \mathbb{E}(u(X_{k+1})v(X_{k+1}))$.

Démonstration.
Soit $(x, y) \in X_{k+1}(\Omega)$ tel que $x \leqslant y$.
La fonction f est $(k+1)$-croissante donc :

$$\forall (x_1, \dots, x_k) \in X_1(\Omega) \times \dots \times X_k(\Omega), \ f(x_1, \dots, x_k, x) \leqslant f(x_1, \dots, x_k, y)$$

d'où, en appliquant cette inégalité à $x_1 = X_1(\omega) \in X_1(\Omega), \dots, x_k = X_k(\omega) \in X_k(\Omega)$, on obtient :

$$\forall \omega \in \Omega, \ f(X_1(\omega), \dots, X_k(\omega), x) \leqslant f(X_1(\omega), \dots, X_k(\omega), y)$$

et enfin :

$$f(X_1, \dots, X_k, x) \leqslant f(X_1, \dots, X_k, y)$$

Comme les variables aléatoires $X_1, \dots, X_k$ sont finies, il suit que les variables aléatoires $f(X_1, \dots, X_k, x)$ et $f(X_1, \dots, X_k, y)$ le sont aussi et donc admettent chacune une espérance.
Par croissance de l'espérance :

$$\mathbb{E}(f(X_1, \dots, X_k, x)) \leqslant \mathbb{E}(f(X_1, \dots, X_k, y))$$

$$\boxed{\text{La fonction } u \text{ est croissante sur } X_{k+1}(\Omega).}$$

Ce raisonnement est également valable pour la fonction g qui est $(k+1)$-croissante.

$$\boxed{\text{Donc } v \text{ est croissante sur } X_{k+1}(\Omega).}$$

On pose $\varphi : x \mapsto u(x)v(x)$ (définie sur $X_{k+1}(\Omega)$).
On applique le théorème de transfert (avec des variables finies) :

$$\begin{aligned}
& \mathbb{E}(u(X_{k+1})v(X_{k+1})) \\
= \ & \mathbb{E}(\varphi(X_{k+1})) \\
= \ & \sum_{x \in X_{k+1}(\Omega)} \varphi(x) \mathbb{P}([X_{k+1} = x]) \\
= \ & \sum_{x \in X_{k+1}(\Omega)} \mathbb{E}(f(X_1, \dots, X_k, x)) \mathbb{E}(g(X_1, \dots, X_k, x)) \mathbb{P}([X_{k+1} = x]) \\
\leqslant \ & \sum_{x \in X_{k+1}(\Omega)} \mathbb{E}(f(X_1, \dots, X_k, x) g(X_1, \dots, X_k, x)) \mathbb{P}([X_{k+1} = x]) && \textit{(d'après la question } \textbf{\textit{10.b)}} \textit{ et car, pour tout } x \in X_{k+1}(\Omega), \mathbb{P}([X_{k+1} = x]) \geqslant 0\textit{)} \\
= \ & \mathbb{E}(Y_{k+1} Z_{k+1}) && \textit{(d'après la question } \textbf{\textit{10.a)}}\textit{)}
\end{aligned}$$

$$\boxed{\text{On a bien : } \mathbb{E}(Y_{k+1} Z_{k+1}) \geqslant \mathbb{E}(u(X_{k+1})v(X_{k+1})).}$$

□

d) En conclure que (H_{k+1}) est vraie. Conclure.

Démonstration.
D'après la question ***10.c)*** : $\mathbb{E}(Y_{k+1}Z_{k+1}) \geqslant \mathbb{E}(u(X_{k+1})v(X_{k+1}))$.
Pour obtenir (H_{k+1}), il reste à montrer :

$$\mathbb{E}(u(X_{k+1})v(X_{k+1})) \geqslant \mathbb{E}(Y_{k+1})\mathbb{E}(Z_{k+1})$$

En effet, on aura alors, par transitivité :

$$\mathbb{E}(Y_{k+1}Z_{k+1}) \geqslant \mathbb{E}(u(X_{k+1})v(X_{k+1})) \geqslant \mathbb{E}(Y_{k+1})\mathbb{E}(Z_{k+1})$$

Or, on sait que u et v sont croissantes sur $X_{k+1}(\Omega)$ donc on peut utiliser la propriété (H_1) :

$$\mathbb{E}(u(X_{k+1})v(X_{k+1})) \geqslant \mathbb{E}(u(X_{k+1}))\mathbb{E}(v(X_{k+1}))$$

Par théorème de transfert :

$$\begin{aligned}\mathbb{E}(u(X_{k+1})) &= \sum_{x\in X_{k+1}(\Omega)} u(x)\mathbb{P}([X_{k+1}=x]) \\ &= \sum_{x\in X_{k+1}(\Omega)} \mathbb{E}(f(X_1,\dots,X_k,x))\mathbb{P}([X_{k+1}=x])\end{aligned}$$

On utilise maintenant le résultat de la question ***10.a)*** dans le cas particulier où f est la fonction ci-dessus et où g est la fonction constante égale à 1 (qui est bien définie et $(k+1)$-croissante sur $X_1(\Omega)\times\cdots\times X_{k+1}(\Omega)$). Dans ce cas particulier : $Z_{k+1}=1$ et $g(X_1,\dots,X_k,x)=1$.
On obtient :

$$\mathbb{E}(Y_{k+1}) = \sum_{x\in X_{k+1}(\Omega)} \mathbb{E}(f(X_1,\dots,X_k,x))\mathbb{P}([X_{k+1}=x])$$

D'où :

$$\mathbb{E}(u(X_{k+1})) = \mathbb{E}(Y_{k+1})$$

Bien entendu, ce résultat est également valable pour v (en prenant cette fois-ci pour f la fonction constante égale à 1) :

$$\mathbb{E}(v(X_{k+1})) = \mathbb{E}(Z_{k+1})$$

Finalement, on a obtenu :

$$\mathbb{E}(Y_{k+1}Z_{k+1}) \geqslant \mathbb{E}(u(X_{k+1})v(X_{k+1})) \geqslant \mathbb{E}(u(X_{k+1}))\mathbb{E}(v(X_{k+1})) = \mathbb{E}(Y_{k+1})\mathbb{E}(Z_{k+1})$$

$$\boxed{\text{La propriété } (H_{k+1}) \text{ est vraie.}}$$

On a fait l'initialisation d'une récurrence lors de la question ***9.*** et on a démontré l'hérédité lors de la question ***10***.

On a donc montré par récurrence : pour tout $k\in\mathbb{N}^*$, (H_k) est vraie. On a démontré l'inégalité de Harris en toute dimension.

□

Commentaire

Les deux questions précédentes cachent une subtilité de notation et d'objet.
La fonction u est définie par : pour tout $x \in X_{k+1}(\Omega), u(x) = \mathbb{E}(f(X_1, \ldots, X_k, x))$.
Ainsi, l'objet $u(X_{k+1})$ est une variable aléatoire, définie par :

$$\forall \omega \in \Omega,\ u(X_{k+1})(\omega) = \mathbb{E}(f(X_1, \ldots, X_k, X_{k+1}(\omega)))$$

Il faut bien se garder d'écrire que $u(X_{k+1}) = \mathbb{E}(f(X_1, \ldots, X_k, X_{k+1}))$ car on changerait alors le sens de l'objet $u(X_{k+1})$ en donnant le même rôle à jouer à $X_1, \ldots, X_k$ et X_{k+1}.
Pour mieux saisir en quoi l'égalité précédente est suspecte, il faut remarquer que $\mathbb{E}(f(X_1, \ldots, X_k, X_{k+1}))$ désigne un nombre réel. Une telle égalité voudrait dire que la variable aléatoire $u(X_{k+1})$ est constante mais il n'y a aucune raison pour que ce soit le cas.

e) La propriété (H_k) reste-t-elle vraie si f et g sont k-décroissantes ? Justifier votre réponse.
Que se passe-t-il si l'une est k-croissante et l'autre k-décroissante ?

Démonstration.
Soient $X_1, \ldots, X_k$ des variables aléatoires finies indépendantes. Soient f et g deux fonctions définies sur $X_1(\Omega) \times \cdots \times X_k(\Omega)$. On pose $Y_k = f(X_1, \ldots, X_k)$ et $Z_k = g(X_1, \ldots, X_k)$.

- Traitons d'abord le cas où f et g sont k-décroissantes.
  Remarquons alors que $-f$ et $-g$ sont k-croissantes et on peut donc appliquer l'inégalité de Harris avec ces fonctions :

$$\begin{aligned}
\mathbb{E}(Y_k Z_k) &= \mathbb{E}(f(X_1, \ldots, X_k) g(X_1, \ldots, X_k)) \\
&= \mathbb{E}\big((-f(X_1, \ldots, X_k))(-g(X_1, \ldots, X_k))\big) \\
&\geqslant \mathbb{E}\big(-f(X_1, \ldots, X_k)\big)\mathbb{E}\big(-g(X_1, \ldots, X_k)\big) && \textit{(d'après l'inégalité de Harris)} \\
&= \big(-\mathbb{E}(f(X_1, \ldots, X_k))\big)\big(-\mathbb{E}(g(X_1, \ldots, X_k))\big) && \textit{(par linéarité de l'espérance)} \\
&= \mathbb{E}(f(X_1, \ldots, X_k))\mathbb{E}(g(X_1, \ldots, X_k)) \\
&= \mathbb{E}(Y_k)\mathbb{E}(Z_k)
\end{aligned}$$

  $\boxed{\text{La propriété } (H_k) \text{ reste vraie si } f \text{ et } g \text{ sont } k\text{-décroissantes.}}$

- Traitons maintenant le cas où f est k-décroissante et g est k-croissante (l'autre cas est analogue).
  D'après l'inégalité de Harris appliquée à $-f$ et g toutes deux k-croissantes :

$$\mathbb{E}\big(-f(X_1, \ldots, X_k) g(X_1, \ldots, X_k)\big) \geqslant \mathbb{E}\big(-f(X_1, \ldots, X_k)\big)\mathbb{E}\big(g(X_1, \ldots, X_k)\big)$$

  Ceci donne, en multipliant par -1 :

$$-\mathbb{E}\big(-f(X_1, \ldots, X_k) g(X_1, \ldots, X_k)\big) \leqslant -\mathbb{E}\big(-f(X_1, \ldots, X_k)\big)\mathbb{E}\big(g(X_1, \ldots, X_k)\big)$$

  Enfin, par linéarité de l'espérance :

$$\mathbb{E}\big(f(X_1, \ldots, X_k) g(X_1, \ldots, X_k)\big) \leqslant \mathbb{E}\big(f(X_1, \ldots, X_k)\big)\mathbb{E}\big(g(X_1, \ldots, X_k)\big)$$

  $\boxed{\text{D'où : } \mathbb{E}(Y_k Z_k) \leqslant \mathbb{E}(Y_k)\mathbb{E}(Z_k). \text{ L'inégalité change de sens dans ce cas.}}$

□

Partie 4 - Inégalité de Janson et application

On reprend les notations de la partie 1.

De plus, pour tout $i \in [\![1, r]\!]$, on pose $Z_{n,i} = \sum\limits_{k=1}^{i} Y_k$. On remarquera que $Z_{n,r} = Z_n$.

Dans cette partie on établit un encadrement de $\mathbb{P}([Z_n = 0])$.

11. Justifier que $\bigcap\limits_{0 \leqslant u < v \leqslant n-1} [T_{u,v} = 0] \subset [Z_n = 0]$. En déduire que $\mathbb{P}([Z_n = 0]) \geqslant (1-p)^{\binom{n}{2}} > 0$.

Démonstration.

- Soit $\omega \in \bigcap\limits_{0 \leqslant u < v \leqslant n-1} [T_{u,v} = 0]$. Alors :

$$\forall\, 0 \leqslant u < v \leqslant n-1,\ T_{u,v}(\omega) = 0$$

Il n'y a donc aucune arête dans le graphe G (il est totalement déconnecté). A fortiori, il n'y a donc aucun triangle (puisqu'un triangle est constitué de trois arêtes). On peut donc conclure que $Z_n(\omega) = 0$, c'est-à-dire : $\omega \in [Z_n = 0]$.

$$\boxed{\text{On a bien : } \bigcap_{0 \leqslant u < v \leqslant n-1} [T_{u,v} = 0] \subset [Z_n = 0].}$$

Commentaire

Rappelons que des événements sont des ensembles. Ainsi, on peut utiliser dans cette question la structure usuelle de démonstration de $A \subset B$, que l'on rappelle ci-dessous.
Soit $x \in A$. Alors ...
...
Donc $x \in B$.

- Par croissance de $\mathbb{P}$, on a :

$$\mathbb{P}([Z_n = 0]) \geqslant \mathbb{P}\left(\bigcap_{0 \leqslant u < v \leqslant n-1} [T_{u,v} = 0] \right)$$

Or :

$$\begin{aligned} \mathbb{P}\left(\bigcap_{0 \leqslant u < v \leqslant n-1} [T_{u,v} = 0] \right) &= \prod_{0 \leqslant u < v \leqslant n-1} \mathbb{P}([T_{u,v} = 0]) && \textit{(par indépendance des variables aléatoires } T_{u,v}\textit{)} \\ &= \prod_{0 \leqslant u < v \leqslant n-1} (1-p) && \textit{(car } T_{u,v} \hookrightarrow \mathcal{B}(p)\textit{)} \\ &= (1-p)^{\binom{n}{2}} \end{aligned}$$

En effet, le nombre de couples (u, v) tels que $0 \leqslant u < v \leqslant n-1$ est égal au nombre de parties à deux éléments de l'ensemble $[\![0, n-1]\!]$, et donc est égal à $\binom{n}{2}$.
De plus, $(1-p)^{\binom{n}{2}} > 0$ car $p < 1$.

$$\boxed{\text{On a bien : } \mathbb{P}([Z_n = 0]) \geqslant (1-p)^{\binom{n}{2}} > 0.}$$

□

12. Montrer que pour tout $i \in [\![1, r]\!]$, $\mathbb{P}([Y_i = 0]) = \mathbb{E}(1 - Y_i)$ et $\mathbb{P}([Z_{n,i} = 0]) = \mathbb{E}\left(\prod_{k=1}^{i}(1 - Y_k)\right)$.

Démonstration.
Soit $i \in [\![1, r]\!]$.

- On sait que $Y_i \hookrightarrow \mathcal{B}\left(p^3\right)$ d'après la question ***2.b)***. On en déduit :

$$\begin{aligned}
\mathbb{P}([Y_i = 0]) &= 1 - \mathbb{P}([Y_i = 1]) \\
&= 1 - p^3 \\
&= 1 - \mathbb{E}(Y_i) \\
&= \mathbb{E}(1 - Y_i) && \textit{(par linéarité de l'espérance)}
\end{aligned}$$

$$\boxed{\text{Pour tout } i \in [\![1, r]\!],\ \mathbb{P}([Y_i = 0]) = \mathbb{E}(1 - Y_i).}$$

- Tout d'abord :

$$\begin{aligned}
\mathbb{P}([Z_{n,i} = 0]) &= \mathbb{P}\left(\left[\sum_{k=1}^{i} Y_k = 0\right]\right) \\
&= \mathbb{P}\left(\bigcap_{k=1}^{i} [Y_k = 0]\right) && \begin{array}{l}\textit{(car une somme de termes positifs est nulle}\\ \textit{si et seulement si tous ses termes sont nuls)}\end{array}
\end{aligned}$$

Par ailleurs, on applique le théorème de transfert d'ordre i à la variable aléatoire $f(Y_1, \ldots, Y_i)$ où $f : (x_1, \ldots, x_i) \mapsto \prod_{k=1}^{i}(1 - x_k)$ est définie sur $Y_1(\Omega) \times \cdots \times Y_i(\Omega) = \{0, 1\}^i$:

$$\mathbb{E}\left(\prod_{k=1}^{i}(1 - Y_k)\right) = \sum_{(x_1,\ldots,x_i)\in\{0,1\}^i} \left(\prod_{k=1}^{i}(1 - x_k)\right) \mathbb{P}\left(\bigcap_{k=1}^{i} [Y_k = x_k]\right)$$

Or :

$$\prod_{k=1}^{i}(1 - x_k) = \begin{cases} 1 & \text{si, pour tout } k \in [\![1, i]\!],\ x_k = 0 \\ 0 & \text{sinon} \end{cases}$$

On en déduit que tous les termes de la somme sont nuls sauf un. Plus précisément, le seul terme non nul de la somme est obtenu pour $(x_1, \ldots, x_i) = (0, \ldots, 0)$. D'où :

$$\mathbb{E}\left(\prod_{k=1}^{i}(1 - Y_k)\right) = \mathbb{P}\left(\bigcap_{k=1}^{i} [Y_k = 0]\right)$$

$$\boxed{\text{Pour tout } i \in [\![1, r]\!],\ \mathbb{P}([Z_{n,i} = 0]) = \mathbb{E}\left(\prod_{k=1}^{i}(1 - Y_k)\right).}$$

□

13. a) On pose $m = \binom{n}{2}$. Justifier brièvement que, pour tout $k \in [\![1, r]\!]$, Y_k s'exprime comme une fonction m-croissante sur $\{0, 1\}^m$ des variables aléatoires $T_{u,v}$ pour $u < v$ éléments de $[\![0, n - 1]\!]$.

En déduire que, pour tout $i \in [\![2, r]\!]$, $1 - Y_i$ puis $\prod_{k=1}^{i-1}(1 - Y_k)$ s'expriment comme des fonctions m-décroissantes des variables aléatoires $T_{u,v}$ pour $u < v$ éléments de $[\![0, n - 1]\!]$.

Démonstration.

- Soit $k \in [\![1,r]\!]$. Posons $t_k = \{x,y,z\}$ avec $x<y<z$.
  D'après la question ***2.a)***, $Y_k = T_{x,y}T_{y,z}T_{x,z}$.
  Numérotons les variables aléatoires $T_{u,v}$ sous la forme $U_1, U_2, \dots, U_m$ de telle sorte que :
  $$U_1 = T_{x,y}, \qquad U_2 = T_{y,z}, \qquad U_3 = T_{x,z}$$
  On pose alors $f : (x_1,\dots,x_m) \mapsto x_1x_2x_3$ définie sur $\{0,1\}^m$ de telle sorte que $Y_k = f(U_1,\dots,U_m)$.
  Montrons que f est m-croissante sur $\{0,1\}^m$.
  - × La fonction f est constante (et donc croissante) par rapport aux variables $x_4,\dots,x_m$. En effet, pour tout $k \in [\![4,m]\!]$, la fonction $x_k \mapsto x_1x_2x_3$ est constante donc croissante.
  - × Pour tout $(x_1,\dots,x_m) \in \{0,1\}^m$, $f(x_1,\dots,x_m) \geqslant 0$ (comme produit de nombres positifs).
  - × Pour tout $(x_1,\dots,x_m) \in \{0,1\}^m$, $f(0,x_1,\dots,x_m) = 0 \leqslant f(1,x_1,\dots,x_m)$, donc f est croissante par rapport à x_1.
  - × On a le même argument pour les variables x_2 et x_3.

La fonction f est m-croissante sur $\{0,1\}^m$ et donc la variable aléatoire Y_k s'exprime comme une fonction m-croissante sur $\{0,1\}^m$ des variables aléatoires $T_{u,v}$ pour $u<v$ éléments de $[\![0,n-1]\!]$.

- Soit $i \in [\![2,r]\!]$. D'après le point précédent, puisque l'opposé d'une fonction croissante est décroissante, la variable aléatoire $-Y_i$ s'exprime comme une fonction m-décroissante sur $\{0,1\}^m$ des variables aléatoires $T_{u,v}$ pour $u<v$ éléments de $[\![0,n-1]\!]$ et il en est de même de $1-Y_i$ (ajouter une constante ne change pas le sens de variations d'une fonction).

Un produit de fonctions décroissantes et positives étant une fonction décroissante, il suit que $\prod\limits_{k=1}^{i-1}(1-Y_k)$ s'exprime également comme une fonction m-décroissante des variables aléatoires $T_{u,v}$ pour $u<v$ éléments de $[\![0,n-1]\!]$.

  □

b) En conclure que, pour tout $i \in [\![2,r]\!]$, $\mathbb{P}([Z_{n,i}=0]) \geqslant \mathbb{P}([Z_{n,i-1}=0])\mathbb{P}([Y_i=0])$ puis que :
$$\mathbb{P}([Z_n=0]) \geqslant \prod_{k=1}^{r} \mathbb{P}([Y_k=0]) \qquad \text{puis} \qquad \mathbb{P}([Z_n=0]) \geqslant (1-p^3)^{\binom{n}{3}}$$

Démonstration.

- Soit $i \in [\![2,r]\!]$.
  $$\begin{aligned}\mathbb{P}([Z_{n,i}=0]) &= \mathbb{E}\left(\prod_{k=1}^{i}(1-Y_k)\right) && \textit{(d'après la question \textbf{12.})}\\ &= \mathbb{E}\left((1-Y_i)\prod_{k=1}^{i-1}(1-Y_k)\right)\end{aligned}$$
  Justifions que l'on peut maintenant appliquer l'inégalité de Harris dans le contexte de la question ***10.e)***, avec deux fonctions m-décroissantes :
  - × Les variables aléatoires $T_{u,v}$ pour $u<v$ éléments de $[\![0,n-1]\!]$ sont finies et indépendantes.
  - × D'après la question ***13.a)***, $1-Y_i$ est une fonction m-décroissante sur $\{0,1\}^m$ des variables aléatoires $T_{u,v}$ pour $u<v$ éléments de $[\![0,n-1]\!]$.
  - × D'après la question ***13.a)***, $\prod\limits_{k=1}^{i-1}(1-Y_k)$ est une fonction m-décroissante sur $\{0,1\}^m$ des variables aléatoires $T_{u,v}$ pour $u<v$ éléments de $[\![0,n-1]\!]$.

Dans ce contexte, l'inégalité de Harris donne :

$$\mathbb{E}\left((1-Y_i)\prod_{k=1}^{i-1}(1-Y_k)\right) \geqslant \mathbb{E}(1-Y_i)\,\mathbb{E}\left(\prod_{k=1}^{i-1}(1-Y_k)\right)$$

Or, d'après la question ***12.*** :

$$\mathbb{E}(1-Y_i) = \mathbb{P}([Y_i=0]) \qquad \text{et} \qquad \mathbb{E}\left(\prod_{k=1}^{i-1}(1-Y_k)\right) = \mathbb{P}([Z_{n,i-1}=0])$$

$$\boxed{\text{On a bien : } \mathbb{P}([Z_{n,i}=0]) \geqslant \mathbb{P}([Z_{n,i-1}=0])\mathbb{P}([Y_i=0]).}$$

- Montrons par récurrence (finie) : $\forall i \in [\![2,r]\!]$, $\mathcal{P}(i)$
  où $\mathcal{P}(i)$: « $\mathbb{P}([Z_{n,i}=0]) \geqslant \prod\limits_{k=1}^{i}\mathbb{P}([Y_k=0])$ ».
  <u>Initialisation</u> :
  D'après le point précédent (en remarquant que $Z_{n,1}=Y_1$) :

$$\mathbb{P}([Z_{n,2}=0]) \geqslant \mathbb{P}([Z_{n,1}=0])\mathbb{P}([Y_2=0]) = \mathbb{P}([Y_1=0])\mathbb{P}([Y_2=0]) = \prod_{k=1}^{2}\mathbb{P}([Y_k=0])$$

D'où $\mathcal{P}(2)$.
<u>Hérédité</u> : soit $i \in [\![2,r-1]\!]$. Supposons $\mathcal{P}(i)$. Montrons $\mathcal{P}(i+1)$.

$$\begin{aligned}\mathbb{P}([Z_{n,i+1}=0]) &\geqslant \mathbb{P}([Z_{n,i}=0])\mathbb{P}([Y_{i+1}=0]) && \textit{(d'après le début de la question, car } i+1\in[\![2,r]\!]\textit{)}\\ &\geqslant \left(\prod_{k=1}^{i}\mathbb{P}([Y_k=0])\right)\mathbb{P}([Y_{i+1}=0]) && \textit{(par hypothèse de récurrence)}\\ &= \prod_{k=1}^{i+1}\mathbb{P}([Y_k=0])\end{aligned}$$

D'où $\mathcal{P}(i+1)$.

On a donc montré par récurrence : pour tout $i\in[\![2,r]\!]$, $\mathbb{P}([Z_{n,i}=0]) \geqslant \prod\limits_{k=1}^{i}\mathbb{P}([Y_k=0])$.

$$\boxed{\text{En particulier pour } i=r : \mathbb{P}([Z_n=0]) = \mathbb{P}([Z_{n,r}=0]) \geqslant \prod_{k=1}^{r}\mathbb{P}([Y_k=0]).}$$

- Pour terminer, rappelons que d'après la question ***2.b)***, pour tout $k\in[\![1,r]\!]$, $Y_k \hookrightarrow \mathcal{B}\left(p^3\right)$.
  On en déduit que :

$$\begin{aligned}\prod_{k=1}^{r}\mathbb{P}([Y_k=0]) &= \prod_{k=1}^{r}(1-p^3)\\ &= (1-p^3)^r\\ &= (1-p^3)^{\binom{n}{3}} && \textit{(d'après la question } \textbf{1.}\textit{)}\end{aligned}$$

$$\boxed{\text{On a bien : } \mathbb{P}([Z_n=0]) \geqslant (1-p^3)^{\binom{n}{3}}.}$$

□

Commentaire

La formule étant donnée dans l'énoncé, elle permet de vérifier le résultat affirmé à la question ***1***, à savoir $r = \dbinom{n}{3}$.

14. *Inégalité de Boole.* Montrer par récurrence sur $k \geqslant 2$ que si $B_1, \ldots, B_k$ sont des événements, on a :

$$\mathbb{P}\left(\bigcup_{i=1}^{k} B_i\right) \leqslant \sum_{i=1}^{k} \mathbb{P}(B_i)$$

Démonstration.
Montrons par récurrence : $\forall k \geqslant 2$, $\mathcal{P}(k)$
où $\mathcal{P}(k)$: « pour tout $B_1, \ldots, B_k$ des événements, $\mathbb{P}\left(\bigcup_{i=1}^{k} B_i\right) \leqslant \sum_{i=1}^{k} \mathbb{P}(B_i)$ ».
Initialisation :
Soient B_1, B_2 deux événements.

$$\begin{aligned} \mathbb{P}(B_1 \cup B_2) &= \mathbb{P}(B_1) + \mathbb{P}(B_2) - \mathbb{P}(B_1 \cap B_2) && \textit{(par formule du crible)} \\ &\leqslant \mathbb{P}(B_1) + \mathbb{P}(B_2) && \textit{(car } \mathbb{P}(B_1 \cap B_2) \geqslant 0\textit{)} \end{aligned}$$

D'où $\mathcal{P}(2)$.
Hérédité : soit $k \geqslant 2$. Supposons $\mathcal{P}(k)$. Montrons $\mathcal{P}(k+1)$.
Soient $B_1, \ldots, B_{k+1}$ des événements. Notons $C_k = \bigcup_{i=1}^{k} B_i$.

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^{k+1} B_i\right) &= \mathbb{P}(C_k \cup B_{k+1}) \\ &= \mathbb{P}(C_k) + \mathbb{P}(B_{k+1}) - \mathbb{P}(C_k \cap B_{k+1}) && \textit{(par formule du crible)} \\ &\leqslant \mathbb{P}(C_k) + \mathbb{P}(B_{k+1}) && \textit{(car } \mathbb{P}(C_k \cap B_{k+1}) \geqslant 0\textit{)} \\ &= \mathbb{P}\left(\bigcup_{i=1}^{k} B_i\right) + \mathbb{P}(B_{k+1}) && \textit{(par définition de } C_k\textit{)} \\ &\leqslant \sum_{i=1}^{k} \mathbb{P}(B_i) + \mathbb{P}(B_{k+1}) && \textit{(par hypothèse de récurrence)} \\ &= \sum_{i=1}^{k+1} \mathbb{P}(B_i) \end{aligned}$$

D'où $\mathcal{P}(k+1)$.

$$\boxed{\text{On a donc montré par récurrence : pour tout } k \geqslant 2, \text{ si } B_1, \ldots, B_k \text{ sont des événements, alors } \mathbb{P}\left(\bigcup_{i=1}^{k} B_i\right) \leqslant \sum_{i=1}^{k} \mathbb{P}(B_i).}$$

□

- Si A est un événement de probabilité non nulle, on rappelle que la probabilité conditionnelle sachant A est notée $\mathbb{P}_A$. On admet qu'elle possède les mêmes propriétés que la probabilité $\mathbb{P}$.
  En particulier l'inégalité de Boole est vérifiée par $\mathbb{P}_A$.
  De plus si X est une variable finie, on note $\mathbb{E}_A(X)$ l'espérance de X pour la probabilité $\mathbb{P}_A$ ce qui signifie que :

$$\mathbb{E}_A(X) = \sum_{x \in X(\Omega)} x \, \mathbb{P}_A([X = x])$$

  Cette espérance conditionnelle possède les mêmes propriétés que l'espérance en particulier l'inégalité de Harris vue dans la partie 3.

15. Soit A, B et C trois événements tels que $\mathbb{P}(B \cap C) \neq 0$ et $\mathbb{P}(A \cap C) \neq 0$.
Montrer que $\mathbb{P}_{B\cap C}(A) \geqslant \mathbb{P}_C(A)\mathbb{P}_{A\cap C}(B)$.

Démonstration.
Tout d'abord, $B \cap C \subset C$, donc, par croissance de $\mathbb{P}$:

$$0 < \mathbb{P}(B \cap C) \leqslant \mathbb{P}(C) \qquad (*)$$

et donc les trois probabilités conditionnelles sont bien définies.
Ensuite :

$$\mathbb{P}_{B\cap C}(A) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(B \cap C)}$$

et

$$\mathbb{P}_C(A)\mathbb{P}_{A\cap C}(B) = \frac{\cancel{\mathbb{P}(A \cap C)}}{\mathbb{P}(C)} \frac{\mathbb{P}(A \cap B \cap C)}{\cancel{\mathbb{P}(A \cap C)}} = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)}$$

D'après $(*)$, par décroissance de la fonction inverse sur $]0, +\infty[$:

$$\frac{1}{\mathbb{P}(B \cap C)} \geqslant \frac{1}{\mathbb{P}(C)}$$

Par positivité de $\mathbb{P}(A \cap B \cap C)$:

$$\frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(B \cap C)} \geqslant \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)}$$

$$\boxed{\text{D'où : } \mathbb{P}_{B\cap C}(A) \geqslant \mathbb{P}_C(A)\mathbb{P}_{A\cap C}(B).}$$ □

► On admet que les probabilités conditionnelles qui interviennent dans la suite sont bien définies.

16. Pour tout $i \in [\![1, r]\!]$, on pose $A_i = [Y_i = 0]$.
On note aussi $I_i = \{j \in [\![1, i-1]\!] \mid j \equiv i\}$ et $J_i = \{j \in [\![1, i-1]\!] \mid j \not\equiv i\}$.
Soit $i \geqslant 2$, on définit $B_i = \bigcap\limits_{j\in I_i} A_j$ et $C_i = \bigcap\limits_{j\in J_i} A_j$, ainsi on a : $B_i \cap C_i = \bigcap\limits_{j=1}^{i-1} A_j$.

a) Justifier que A_i et C_i sont indépendants. En déduire que $\mathbb{P}_{B_i\cap C_i}(\overline{A_i}) \geqslant \mathbb{P}(\overline{A_i})\mathbb{P}_{\overline{A_i}\cap C_i}(B_i)$.

Démonstration.
- Soit $j \in J_i$. On a alors $j \neq i$ et $j \not\equiv i$ et donc $(i, j) \in \mathcal{F}$ où $\mathcal{F}$ est l'ensemble défini avant la question ***3.a)***.
  Rappelons que, d'après la question ***3.b)***, les variables aléatoires Y_i et Y_j sont indépendantes car on peut les écrire :

  $$Y_i = T_{u_i,v_i}T_{v_i,w_i}T_{u_i,w_i} \qquad \text{et} \qquad Y_j = T_{u_j,v_j}T_{v_j,w_j}T_{u_j,w_j}$$

  où $u_i < v_i < w_i$ et $u_j < v_j < w_j$ et où ces six variables aléatoires $T_{u,v}$ sont toutes distinctes.
  Numérotons comme en question ***13.a)*** les variables aléatoires $T_{u,v}$ sous la forme $U_1, U_2, \dots, U_m$ (où $m = \binom{n}{2}$) de telle sorte que :

  $$Y_i = U_1U_2U_3$$

  × D'une part :

  $$A_i = [U_1U_2U_3 = 0]$$

× D'autre part :

$$C_i = \bigcap_{j\in J_i} A_j = \bigcap_{j\in J_i} [Y_j = 0] = \left[\sum_{j\in J_i} Y_j = 0\right]$$ *(car une somme de termes positifs est nulle si et seulement si tous ses termes sont nuls)*

et, d'après la rappel fait ci-dessus, il existe une fonction f telle que $\sum\limits_{j\in J_i} Y_j = f(U_4, \ldots, U_m)$.

On sait que les variables aléatoires $U_1, \ldots, U_m$ sont mutuellement indépendantes. D'après le lemme des coalitions, les variables aléatoires $U_1U_2U_3$ et $f(U_4, \ldots, U_m)$ sont indépendantes.

$$\boxed{\text{On en déduit que } A_i \text{ et } C_i \text{ sont indépendants.}}$$

- A_i et C_i sont indépendants donc $\overline{A_i}$ et C_i sont indépendants.
- D'après la question ***15.*** :

$$\mathbb{P}_{B_i\cap C_i}(\overline{A_i}) \geqslant \mathbb{P}_{C_i}(\overline{A_i})\mathbb{P}_{\overline{A_i}\cap C_i}(B_i)$$

  et par indépendance de $\overline{A_i}$ et C_i :

$$\mathbb{P}_{C_i}(\overline{A_i}) = \mathbb{P}(\overline{A_i})$$

$$\boxed{\text{On a bien : } \mathbb{P}_{B_i\cap C_i}(\overline{A_i}) \geqslant \mathbb{P}(\overline{A_i})\mathbb{P}_{\overline{A_i}\cap C_i}(B_i).}$$

□

Commentaire

Si $J_i = \varnothing$, il n'existe pas d'élement j de J_i pour faire la preuve, mais dans ce cas $C_i = \Omega$ et donc A_i et C_i sont indépendants. En effet :

$$\mathbb{P}(A_i \cap C_i) = \mathbb{P}(A_i \cap \Omega) = \mathbb{P}(A_i) = \mathbb{P}(A_i) \times 1 = \mathbb{P}(A_i)\mathbb{P}(C_i)$$

b) Établir que $\mathbb{P}_{\overline{A_i}\cap C_i}(B_i) \geqslant 1 - \sum\limits_{j\in I_i} \mathbb{P}_{\overline{A_i}\cap C_i}(\overline{A_j})$.

Démonstration.
Tout d'abord, $\mathbb{P}_{\overline{A_i}\cap C_i}$ étant une (application) probabilité :

$$\mathbb{P}_{\overline{A_i}\cap C_i}(B_i) = 1 - \mathbb{P}_{\overline{A_i}\cap C_i}(\overline{B_i})$$

Or : $\overline{B_i} = \overline{\bigcap\limits_{j\in I_i} A_j} = \bigcup\limits_{j\in I_i} \overline{A_j}$. D'où :

$$\mathbb{P}_{\overline{A_i}\cap C_i}(B_i) = 1 - \mathbb{P}_{\overline{A_i}\cap C_i}\left(\bigcup_{j\in I_i} \overline{A_j}\right)$$

D'après l'inégalité de Boole (l'ensemble I_i étant fini) :

$$\mathbb{P}_{\overline{A_i}\cap C_i}\left(\bigcup_{j\in I_i} \overline{A_j}\right) \leqslant \sum_{j\in I_i} \mathbb{P}_{\overline{A_i}\cap C_i}(\overline{A_j})$$

$$\boxed{\text{On a bien : } \mathbb{P}_{\overline{A_i}\cap C_i}(B_i) \geqslant 1 - \sum_{j\in I_i} \mathbb{P}_{\overline{A_i}\cap C_i}(\overline{A_j}).}$$

□

c) **On admet provisoirement que pour** $j \in [\![1, i-1]\!]$:

$$\mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) \leqslant \mathbb{P}_{\overline{A_i}}(\overline{A_j}) \qquad (1)$$

En déduire que, $\mathbb{P}_{B_i \cap C_i}(A_i) \leqslant 1 - \mathbb{P}(\overline{A_i}) \left(1 - \sum\limits_{j \in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})\right)$.

Démonstration.

- Tout d'abord, d'après la question ***16.a)*** :

$$\mathbb{P}_{B_i \cap C_i}(A_i) = 1 - \mathbb{P}_{B_i \cap C_i}(\overline{A_i}) \leqslant 1 - \mathbb{P}(\overline{A_i}) \mathbb{P}_{\overline{A_i} \cap C_i}(B_i)$$

- Ensuite, d'après la question ***16.b)*** :

$$\mathbb{P}_{\overline{A_i} \cap C_i}(B_i) \geqslant 1 - \sum_{j \in I_i} \mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j})$$

  Et, puisque $-\mathbb{P}(\overline{A_i}) \leqslant 0$:

$$1 - \mathbb{P}(\overline{A_i}) \mathbb{P}_{\overline{A_i} \cap C_i}(B_i) \leqslant 1 - \mathbb{P}(\overline{A_i}) \left(1 - \sum_{j \in I_i} \mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j})\right)$$

- Enfin, d'après (1), pour tout $j \in [\![1, i-1]\!]$:

$$\mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) \leqslant \mathbb{P}_{\overline{A_i}}(\overline{A_j})$$

  En sommant ces inégalités pour j variant dans $I_i \subset [\![1, i-1]\!]$, on obtient :

$$\sum_{j \in I_i} \mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) \leqslant \sum_{j \in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})$$

  D'où :

$$1 - \sum_{j \in I_i} \mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) \geqslant 1 - \sum_{j \in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})$$

  Et, puisque $-\mathbb{P}(\overline{A_i}) \leqslant 0$:

$$1 - \mathbb{P}(\overline{A_i}) \left(1 - \sum_{j \in I_i} \mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j})\right) \leqslant 1 - \mathbb{P}(\overline{A_i}) \left(1 - \sum_{j \in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})\right)$$

$$\boxed{\text{Par transitivité, on a bien : } \mathbb{P}_{B_i \cap C_i}(A_i) \leqslant 1 - \mathbb{P}(\overline{A_i}) \left(1 - \sum_{j \in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})\right).}$$

□

d) Justifier que pour tout $x \in \mathbb{R}$, $1 - x \leqslant \exp(-x)$ et en déduire que :

$$\mathbb{P}_{B_i \cap C_i}(A_i) \leqslant \exp\left(-\mathbb{P}(\overline{A_i}) + \sum_{j \in I_i} \mathbb{P}(\overline{A_j} \cap \overline{A_i})\right)$$

Démonstration.

- La fonction $h = \exp$ est convexe sur $\mathbb{R}$, son graphe est donc au dessus de chacune de ses tangentes, en particulier celle en 0 admettant pour équation : $y = h(0) + h'(0)(x - 0) = 1 + x$. On en déduit que, pour tout $u \in \mathbb{R}$, $\mathrm{e}^u \geqslant 1 + u$ et donc, en posant $u = -x \in \mathbb{R}$:

$$\boxed{\text{pour tout } x \in \mathbb{R},\ 1 - x \leqslant \exp(-x).}$$

- Ensuite :

$$\mathbb{P}_{B_i\cap C_i}(A_i) \leqslant 1 - \mathbb{P}(\overline{A_i})\left(1 - \sum_{j\in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})\right) \qquad \textit{(d'après la question } \textbf{\textit{16.c)}}\textit{)}$$

$$\leqslant \exp\left(-\mathbb{P}(\overline{A_i})\left(1 - \sum_{j\in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})\right)\right) \qquad \textit{(d'après l'inégalité de convexité appliquée à } x = \mathbb{P}(\overline{A_i})\left(1 - \sum_{j\in I_i} \mathbb{P}_{\overline{A_i}}(\overline{A_j})\right)\textit{)}$$

$$= \exp\left(-\mathbb{P}(\overline{A_i}) + \sum_{j\in I_i} \mathbb{P}(\overline{A_i})\mathbb{P}_{\overline{A_i}}(\overline{A_j})\right)$$

$$= \exp\left(-\mathbb{P}(\overline{A_i}) + \sum_{j\in I_i} \mathbb{P}(\overline{A_j}\cap\overline{A_i})\right)$$

$$\boxed{\text{On a donc bien : } \mathbb{P}_{B_i\cap C_i}(A_i) \leqslant \exp\left(-\mathbb{P}(\overline{A_i}) + \sum_{j\in I_i} \mathbb{P}(\overline{A_j}\cap\overline{A_i})\right).}$$

□

17. On rappelle que $\Delta_n = \sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_iY_j)$ où $\mathcal{E}$ a été défini dans la partie 1 à la suite de la question ***2***.

a) Montrer que $\mathbb{P}([Z_n = 0]) = \mathbb{P}\left(\bigcap_{i=1}^r A_i\right) = \mathbb{P}(A_1)\prod_{i=2}^r \mathbb{P}_{B_i\cap C_i}(A_i)$.

Démonstration.

- On a déjà montré en question ***12*** que pour tout $i \in [\![1, r]\!]$,

$$\mathbb{P}([Z_{n,i} = 0]) = \mathbb{P}\left(\bigcap_{k=1}^{i} [Y_k = 0]\right)$$

(car une somme de termes positifs est nulle si et seulement si tous ses termes sont nuls).

$$\boxed{\text{En particulier pour } i = r\text{, on a } \mathbb{P}([Z_n = 0]) = \mathbb{P}\left(\bigcap_{i=1}^r [Y_i = 0]\right) = \mathbb{P}\left(\bigcap_{i=1}^r A_i\right).}$$

- D'après la formule des probabilités composées :

$$\mathbb{P}\left(\bigcap_{i=1}^r A_i\right) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \cdots \times \mathbb{P}_{A_1\cap\cdots\cap A_{r-1}}(A_r)$$

Or, d'après l'énoncé de la question ***16***, pour tout $i \in [\![2, r]\!]$, $B_i \cap C_i = \bigcap_{j=1}^{i-1} A_j$. D'où :

$$\mathbb{P}_{A_1}(A_2) \times \cdots \times \mathbb{P}_{A_1\cap\cdots\cap A_{r-1}}(A_r) = \mathbb{P}_{B_2\cap C_2}(A_2) \times \cdots \times \mathbb{P}_{B_r\cap C_r}(A_r) = \prod_{i=2}^r \mathbb{P}_{B_i\cap C_i}(A_i)$$

$$\boxed{\text{On obtient bien : } \mathbb{P}\left(\bigcap_{i=1}^r A_i\right) = \mathbb{P}(A_1)\prod_{i=2}^r \mathbb{P}_{B_i\cap C_i}(A_i).}$$

□

b) En conclure que :

$$\mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\mathbb{E}(Z_n) + \frac{\Delta_n}{2}\right) \quad \text{(inégalité de Janson)}$$

Démonstration.

- Tout d'abord :

$$\begin{aligned}
\mathbb{P}([Z_n = 0]) &= \mathbb{P}(A_1) \prod_{i=2}^{r} \mathbb{P}_{B_i \cap C_i}(A_i) && \textit{(d'après la question } \textbf{\textit{17.a)}}\textit{)} \\
&\leqslant \mathbb{P}(A_1) \prod_{i=2}^{r} \exp\left(-\mathbb{P}(\overline{A_i}) + \sum_{j \in I_i} \mathbb{P}(\overline{A_j} \cap \overline{A_i})\right) && \textit{(d'après la question } \textbf{\textit{16.d)}}\textit{)} \\
&= \mathbb{P}(A_1) \exp\left(-\sum_{i=2}^{r} \mathbb{P}(\overline{A_i}) + \sum_{i=2}^{r} \sum_{j \in I_i} \mathbb{P}(\overline{A_j} \cap \overline{A_i})\right) && \textit{(par propriété de morphisme de } \exp\textit{)}
\end{aligned}$$

  De plus, par inégalité de convexité (question ***16.d)***) :

$$\mathbb{P}(A_1) = 1 - \mathbb{P}(\overline{A_1}) \leqslant \exp\left(-\mathbb{P}(\overline{A_1})\right)$$

  D'où :

$$\begin{aligned}
\mathbb{P}([Z_n = 0]) &\leqslant \exp\left(-\mathbb{P}(\overline{A_1})\right) \exp\left(-\sum_{i=2}^{r} \mathbb{P}(\overline{A_i}) + \sum_{i=2}^{r} \sum_{j \in I_i} \mathbb{P}(\overline{A_j} \cap \overline{A_i})\right) \\
&\leqslant \exp\left(-\sum_{i=1}^{r} \mathbb{P}(\overline{A_i}) + \sum_{i=2}^{r} \sum_{j \in I_i} \mathbb{P}(\overline{A_j} \cap \overline{A_i})\right)
\end{aligned}$$

- Ensuite :

$$\begin{aligned}
\mathbb{E}(Z_n) &= \mathbb{E}\left(\sum_{i=1}^{r} Y_i\right) \\
&= \sum_{i=1}^{r} \mathbb{E}(Y_i) && \textit{(par linéarité de l'espérance)} \\
&= \sum_{i=1}^{r} \mathbb{P}([Y_i = 1]) && \textit{(car, pour tout } i \in [\![1, r]\!]\textit{, } Y_i \textit{ suit une loi de Bernoulli)} \\
&= \sum_{i=1}^{r} \mathbb{P}(\overline{A_i}) && \textit{(car, pour tout } i \in [\![1, r]\!]\textit{, } A_i = [Y_i = 0]\textit{)}
\end{aligned}$$

- Fixons momentanément $i \in [\![2, r]\!]$ et $j \in I_i$. Remarquons que $Y_i Y_j$ est une variable de Bernoulli (comme produit de variables de Bernoulli) et donc

$$\mathbb{E}(Y_i Y_j) = \mathbb{P}([Y_i Y_j = 1]) = \mathbb{P}([Y_i = 1] \cap [Y_j = 1]) = \mathbb{P}(\overline{A_i} \cap \overline{A_j})$$

- On peut alors conclure :

$$\mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\mathbb{E}(Z_n) + \sum_{i=2}^{r} \sum_{j \in I_i} \mathbb{E}(Y_i Y_j)\right)$$

  Il reste à montrer :

$$\sum_{i=2}^{r} \sum_{j \in I_i} \mathbb{E}(Y_i Y_j) = \frac{\Delta_n}{2}$$

Or :

$$\begin{aligned}
\Delta_n &= \sum_{(i,j)\in\mathcal{E}} \mathbb{E}(Y_i Y_j) \\
&= 2 \sum_{\substack{(i,j)\in\mathcal{E} \\ j<i}} \mathbb{E}(Y_i Y_j) && \textit{(chaque paire } \{i,j\} \textit{ apparaît deux fois : une fois quand } i<j\textit{, une fois quand } j<i\textit{)} \\
&= 2 \sum_{\substack{i\equiv j \\ j<i}} \mathbb{E}(Y_i Y_j) \\
&= 2 \sum_{i=2}^{r} \sum_{\substack{j=1 \\ j\equiv i}}^{i-1} \mathbb{E}(Y_i Y_j) \\
&= 2 \sum_{i=2}^{r} \sum_{j\in I_i} \mathbb{E}(Y_i Y_j)
\end{aligned}$$

$$\boxed{\text{On obtient bien : } \mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\mathbb{E}(Z_n) + \frac{\Delta_n}{2}\right).}$$

□

Commentaire

Cette question est à considérer comme très difficile car elle nécessite de prendre plusieurs initiatives. Il faut réussir à s'en rendre compte et ne pas passer trop de temps dessus si on bloque car elle ne rapportera pas assez de points pour le justifier. Cependant, toute résolution partielle sera récompensée.

c) En déduire l'encadrement :

$$(1-p^3)^{\binom{n}{3}} \leqslant \mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\binom{n}{3}p^3 + \frac{a_n}{2}p^5\right)$$

Démonstration.

- D'après la question ***13.b)*** :
$$(1-p^3)^{\binom{n}{3}} \leqslant \mathbb{P}([Z_n = 0])$$
- Démontrons maintenant l'autre inégalité :

$$\begin{aligned}
\mathbb{P}([Z_n = 0]) &\leqslant \exp\left(-\mathbb{E}(Z_n) + \frac{\Delta_n}{2}\right) && \textit{(d'après l'inégalité de Janson de } \textbf{\textit{17.b)}}\textit{)} \\
&\leqslant \exp\left(-\binom{n}{3}p^3 + \frac{a_n}{2}p^5\right) && \textit{(d'après les questions } \textbf{\textit{2.c)}} \textit{ et } \textbf{\textit{4.}}\textit{)}
\end{aligned}$$

$$\boxed{\begin{array}{c}\text{On a bien l'encadrement voulu :} \\ (1-p^3)^{\binom{n}{3}} \leqslant \mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\binom{n}{3}p^3 + \frac{a_n}{2}p^5\right).\end{array}}$$

□

18. Soit c un réel strictement positif.

a) Montrer que $\lim\limits_{n\to+\infty} \ -\binom{n}{3}\left(\frac{c}{n}\right)^3 + \frac{a_n}{2}\left(\frac{c}{n}\right)^5 = -\frac{c^3}{6}$.

Démonstration.

- D'une part :
$$\binom{n}{3} = \frac{n(n-1)(n-2)}{3!} \underset{n\to+\infty}{\sim} \frac{n^3}{6}$$
On en déduit :
$$\binom{n}{3}\left(\frac{c}{n}\right)^3 \underset{n\to+\infty}{\sim} \frac{n^3}{6}\frac{c^3}{n^3} = \frac{c^3}{6}$$
et donc :
$$\lim_{n\to+\infty} \ -\binom{n}{3}\left(\frac{c}{n}\right)^3 = -\frac{c^3}{6}$$
- D'autre part, d'après la question ***5.b)*** :
$$a_n = \frac{n(n-1)(n-2)(n-3)}{2} \underset{n\to+\infty}{\sim} \frac{n^4}{2}$$
On en déduit :
$$\frac{a_n}{2}\left(\frac{c}{n}\right)^5 \underset{n\to+\infty}{\sim} \frac{n^4}{4}\frac{c^5}{n^5} = \frac{c^5}{4n}$$
et donc :
$$\lim_{n\to+\infty} \frac{a_n}{2}\left(\frac{c}{n}\right)^5 = 0$$

$$\boxed{\text{D'où : } \lim_{n\to+\infty} \ -\binom{n}{3}\left(\frac{c}{n}\right)^3 + \frac{a_n}{2}\left(\frac{c}{n}\right)^5 = -\frac{c^3}{6}.}$$

□

b) Établir que $\lim\limits_{n\to+\infty} \binom{n}{3}\ln\left(1-\frac{c^3}{n^3}\right) = -\frac{c^3}{6}$.

Démonstration.

Puisque $\frac{c^3}{n^3} \underset{n\to+\infty}{\longrightarrow} 0$, on a :
$$\ln\left(1-\frac{c^3}{n^3}\right) \underset{n\to+\infty}{\sim} -\frac{c^3}{n^3}$$
On en déduit :
$$\binom{n}{3}\ln\left(1-\frac{c^3}{n^3}\right) \underset{n\to+\infty}{\sim} -\frac{n^3}{6}\frac{c^3}{n^3} = -\frac{c^3}{6}$$
et donc
$$\boxed{\lim_{n\to+\infty} \binom{n}{3}\ln\left(1-\frac{c^3}{n^3}\right) = -\frac{c^3}{6}.}$$

□

c) On suppose que $n > c$ et $p = \dfrac{c}{n}$. En déduire la limite de $\mathbb{P}([Z_n = 0])$ quand $n \to +\infty$.

Démonstration.
D'après la question ***17.c)*** :

$$(1-p^3)^{\binom{n}{3}} \leqslant \mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\binom{n}{3}p^3 + \frac{a_n}{2}p^5\right)$$

Donc :

$$\left(1-\left(\frac{c}{n}\right)^3\right)^{\binom{n}{3}} \leqslant \mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\binom{n}{3}\left(\frac{c}{n}\right)^3 + \frac{a_n}{2}\left(\frac{c}{n}\right)^5\right)$$

D'où :

$$\exp\left(\binom{n}{3}\ln\left(1-\frac{c^3}{n^3}\right)\right) \leqslant \mathbb{P}([Z_n = 0]) \leqslant \exp\left(-\binom{n}{3}\left(\frac{c}{n}\right)^3 + \frac{a_n}{2}\left(\frac{c}{n}\right)^5\right)$$

× D'après la question ***18.a)*** : $\lim\limits_{n\to+\infty} \; -\binom{n}{3}\left(\frac{c}{n}\right)^3 + \frac{a_n}{2}\left(\frac{c}{n}\right)^5 = -\frac{c^3}{6}$.

On en déduit, par continuité de l'exponentielle en $-\dfrac{c^3}{6}$:

$$\lim_{n\to+\infty} \exp\left(-\binom{n}{3}\left(\frac{c}{n}\right)^3 + \frac{a_n}{2}\left(\frac{c}{n}\right)^5\right) = \exp\left(-\frac{c^3}{6}\right)$$

× D'après la question ***18.b)*** : $\lim\limits_{n\to+\infty} \binom{n}{3}\ln\left(1-\frac{c^3}{n^3}\right) = -\frac{c^3}{6}$.

On en déduit, par continuité de l'exponentielle en $-\dfrac{c^3}{6}$:

$$\lim_{n\to+\infty} \exp\left(\binom{n}{3}\ln\left(1-\frac{c^3}{n^3}\right)\right) = \exp\left(-\frac{c^3}{6}\right)$$

$$\boxed{\text{Par théorème d'encadrement : } \lim_{n\to+\infty} \mathbb{P}([Z_n = 0]) = \exp\left(-\frac{c^3}{6}\right).}$$

□

Commentaire

On peut traiter la question ***18*** en admettant les résultats précédents même si l'on ne comprend pas bien où va le sujet. Il est bon de repérer en amont les questions qui sont largement indépendantes de la compréhension du problème et il faut garder du temps pour elles.

19. On reprend les notations de la partie 2. L'exécution de l'instruction `fonctionMystere(100)` affiche dans la console **Python** `0.849`. Est-ce cohérent avec le résultat de la question précédente si on considère que pour x assez petit, e^{-x} est proche de $1 - x + \dfrac{x^2}{2}$?

Démonstration.
Lors de l'exécution de l'instruction `fonctionMystere(100)`, les graphes sont générés aléatoirement avec $p = \dfrac{1}{n}$. Cela correspond donc au cas où $c = np = 1$. On sait d'après la question ***8*** que le nombre renvoyé par `fonctionMystere(100)` doit être une approximation de $\mathbb{P}([Z_{100} = 0])$.

Comparons donc $0,849$ et $\lim\limits_{n\to+\infty} \mathbb{P}([Z_n = 0]) = \exp\left(-\dfrac{c^3}{6}\right) = \exp\left(-\dfrac{1}{6}\right)$.

D'après l'énoncé, e^{-x} est proche de $1 - x + \dfrac{x^2}{2}$ pour x assez petit. On considérera dans la suite que $\dfrac{1}{6}$ est assez petit pour que cette approximation soit valide. On obtient :

$$\exp\left(-\frac{1}{6}\right) \simeq 1 - \frac{1}{6} + \frac{\left(\frac{1}{6}\right)^2}{2} = \frac{5}{6} + \frac{1}{2 \times 6^2} = \frac{61}{72} = 0,847\ldots$$

(les premières décimales sont calculées par divisions euclidiennes successives).

$$\boxed{\text{Ainsi, les deux résultats sont cohérents.}}$$

□

20. *Démonstration de (1)* - Soit m un entier plus grand que 2. On considère $X_1, \ldots, X_m$ des variables de Bernoulli indépendantes et I un sous ensemble de $[\![1, m]\!]$. On note J le complémentaire de I dans $[\![1, m]\!]$. On note A l'événement $\left[\prod_{i \in I} X_i = 1\right]$.

a) Montrer que, pour tout $(x_1, \ldots, x_m) \in \{0, 1\}^m$,

$$\mathbb{P}_A([X_1 = x_1] \cap \cdots \cap [X_m = x_m]) = \begin{cases} \prod_{i \in J} \mathbb{P}([X_i = x_i]) & \text{si } \prod_{i \in I} x_i = 1 \\ 0 & \text{sinon.} \end{cases}$$

Démonstration.
Soit $(x_1, \ldots, x_m) \in \{0, 1\}^m$. Deux cas se présentent :

× Premier cas : pour tout $k \in I$, $x_k = 1$.
Alors $\prod_{i \in I} x_i = 1$.

× Deuxième cas : il existe $k \in I$ tel que $x_k = 0$.
Alors $\prod_{i \in I} x_i = 0$.

On obtient donc l'équivalence :

$$\prod_{i \in I} x_i = 1 \iff \forall k \in I, x_k = 1$$

Cette équivalence permet d'affirmer que : $A = \left[\prod_{i \in I} X_i = 1\right] = \bigcap_{i \in I} [X_i = 1]$.

On en déduit :

$$\begin{aligned} \mathbb{P}_A([X_1 = x_1] \cap \cdots \cap [X_m = x_m]) &= \frac{\mathbb{P}(A \cap [X_1 = x_1] \cap \cdots \cap [X_m = x_m])}{\mathbb{P}(A)} \\ &= \frac{\mathbb{P}\left(\left(\bigcap_{i \in I} [X_i = 1]\right) \cap \left(\bigcap_{k=1}^{m} [X_k = x_k]\right)\right)}{\mathbb{P}\left(\bigcap_{i \in I} [X_i = 1]\right)} \end{aligned}$$

Distinguons deux cas :

× Premier cas : $\prod_{i \in I} x_i = 1$.

Alors pour tout $k \in I$, $x_k = 1$, et donc

$$\begin{aligned}\bigcap_{k=1}^{m} [X_k = x_k] &= \Big(\bigcap_{k \in I} [X_k = x_k]\Big) \cap \Big(\bigcap_{k \notin I} [X_k = x_k]\Big) \\ &= \Big(\bigcap_{k \in I} [X_k = 1]\Big) \cap \Big(\bigcap_{k \notin I} [X_k = x_k]\Big) \\ &\subset \bigcap_{k \in I} [X_k = 1]\end{aligned}$$

D'où :

$$\Big(\bigcap_{i \in I} [X_i = 1]\Big) \cap \Big(\bigcap_{k=1}^{m} [X_k = x_k]\Big) = \bigcap_{k=1}^{m} [X_k = x_k]$$

Ainsi :

$$\begin{aligned}\mathbb{P}_A([X_1 = x_1] \cap \cdots \cap [X_m = x_m]) &= \frac{\mathbb{P}\Big(\bigcap_{k=1}^{m} [X_k = x_k]\Big)}{\mathbb{P}\Big(\bigcap_{i \in I} [X_i = 1]\Big)} \\ &= \frac{\prod_{k=1}^{m} \mathbb{P}([X_k = x_k])}{\prod_{i \in I} \mathbb{P}([X_i = 1])} && \textit{(par indépendance des variables aléatoires } X_1, \dots, X_m\textit{)} \\ &= \frac{\prod_{k=1}^{m} \mathbb{P}([X_k = x_k])}{\prod_{k \in I} \mathbb{P}([X_k = x_k])} && \textit{(car pour tout } k \in I,\ x_k = 1\textit{)} \\ &= \prod_{\substack{k=1 \\ k \notin I}}^{m} \mathbb{P}([X_k = x_k]) \\ &= \prod_{i \in J} \mathbb{P}([X_i = x_i]) && \textit{(car } I \textit{ et } J \textit{ sont complémentaires dans } [\![1, m]\!]\textit{)}\end{aligned}$$

× Deuxième cas : $\prod_{i \in I} x_i = 0$.

Alors il existe $k \in I$ tel que $x_k = 0$.

– D'une part :

$$\bigcap_{i \in I} [X_i = 1] = \left(\bigcap_{\substack{i \in I \\ i \neq k}} [X_i = 1]\right) \cap [X_k = 1] \subset [X_k = 1]$$

– D'autre part :

$$\bigcap_{i=1}^{m} [X_i = x_i] = \left(\bigcap_{\substack{i=1 \\ i \neq k}}^{m} [X_i = x_i]\right) \cap [X_k = x_k] = \left(\bigcap_{\substack{i=1 \\ i \neq k}}^{m} [X_i = x_i]\right) \cap [X_k = 0] \subset [X_k = 0]$$

On en déduit :

$$\left(\bigcap_{i\in I}[X_i=1]\right)\cap\left(\bigcap_{i=1}^{m}[X_i=x_i]\right)\subset[X_k=1]\cap[X_k=0]=\varnothing$$

ce qui permet de conclure :

$$\left(\bigcap_{i\in I}[X_i=1]\right)\cap\left(\bigcap_{i=1}^{m}[X_i=x_i]\right)=\varnothing$$

puis :

$$\mathbb{P}_A([X_1=x_1]\cap\cdots\cap[X_m=x_m])=0$$

$$\boxed{\text{On a montré : } \mathbb{P}_A([X_1=x_1]\cap\cdots\cap[X_m=x_m])=\begin{cases}\prod\limits_{i\in J}\mathbb{P}([X_i=x_i]) & \text{si } \prod\limits_{i\in I}x_i=1\\ 0 & \text{sinon.}\end{cases}}$$ □

b) En déduire que les variables aléatoires $X_1,\ldots,X_m$ sont indépendantes pour la probabilité conditionnelle $\mathbb{P}_A$.

Démonstration.
Soit $(x_1,\ldots,x_m)\in\{0,1\}^m$. Il s'agit de montrer :

$$\mathbb{P}_A\left(\bigcap_{k=1}^{m}[X_k=x_k]\right)=\prod_{k=1}^{m}\mathbb{P}_A([X_k=x_k])$$

Or, d'après la question précédente :

$$\mathbb{P}_A\left(\bigcap_{k=1}^{m}[X_k=x_k]\right)=\begin{cases}\prod\limits_{i\in J}\mathbb{P}([X_i=x_i]) & \text{si } \prod\limits_{i\in I}x_i=1\\ 0 & \text{sinon.}\end{cases}$$

On souhaite donc montrer :

$$\prod_{k=1}^{m}\mathbb{P}_A([X_k=x_k])=\begin{cases}\prod\limits_{i\in J}\mathbb{P}([X_i=x_i]) & \text{si } \prod\limits_{i\in I}x_i=1\\ 0 & \text{sinon.}\end{cases}$$

Deux cas se présentent :

× Premier cas : $\prod\limits_{i\in I}x_i=1$.
Alors pour tout $k\in I$, $x_k=1$. On en déduit, pour tout $k\in I$:

$$\begin{aligned}A&=\bigcap_{i\in I}[X_i=1]\\&=\left(\bigcap_{\substack{i\in I\\ i\neq k}}[X_i=1]\right)\cap[X_k=1]\\&\subset[X_k=1]\\&=[X_k=x_k]\end{aligned}$$

et donc :

$$\mathbb{P}_A([X_k = x_k]) = 1$$

Ainsi :

$$\prod_{k=1}^{m} \mathbb{P}_A([X_k = x_k]) = \prod_{k\in I} \mathbb{P}_A([X_k = x_k]) \times \prod_{k\in J} \mathbb{P}_A([X_k = x_k]) = \prod_{k\in J} \mathbb{P}_A([X_k = x_k])$$

× Deuxième cas : $\displaystyle\prod_{i\in I} x_i = 0$.

Alors il existe $k \in I$ tel que $x_k = 0$. On en déduit :

$$\begin{aligned}
A \cap [X_k = x_k] &= A \cap [X_k = 0] \\
&= \left(\bigcap_{i\in I} [X_i = 1]\right) \cap [X_k = 0] \\
&= \left(\left(\bigcap_{\substack{i\in I\\ i\neq k}} [X_i = 1]\right) \cap [X_k = 1]\right) \cap [X_k = 0] \\
&= \left(\bigcap_{\substack{i\in I\\ i\neq k}} [X_i = 1]\right) \cap ([X_k = 1] \cap [X_k = 0]) \\
&= \left(\bigcap_{\substack{i\in I\\ i\neq k}} [X_i = 1]\right) \cap \varnothing \\
&= \varnothing
\end{aligned}$$

D'où :

$$\mathbb{P}_A([X_k = x_k]) = \frac{\mathbb{P}(A \cap [X_k = x_k])}{\mathbb{P}(A)} = \frac{\mathbb{P}(\varnothing)}{\mathbb{P}(A)} = 0$$

On en déduit :

$$\prod_{i=1}^{m} \mathbb{P}_A([X_i = x_i]) = \mathbb{P}_A([X_k = x_k]) \times \prod_{\substack{i=1\\ i\neq k}}^{m} \mathbb{P}_A([X_i = x_i]) = 0 \times \prod_{\substack{i=1\\ i\neq k}}^{m} \mathbb{P}_A([X_i = x_i]) = 0$$

Finalement :

$$\prod_{k=1}^{m} \mathbb{P}_A([X_k = x_k]) = \begin{cases} \displaystyle\prod_{i\in J} \mathbb{P}([X_i = x_i]) & \text{si } \displaystyle\prod_{i\in I} x_i = 1 \\ 0 & \text{sinon.} \end{cases}$$

$$\boxed{\text{Les variables aléatoires } X_1, \dots, X_m \text{ sont indépendantes pour la probabilité conditionnelle } \mathbb{P}_A.}$$

□

▶ Soit $i \geqslant 2$, on reprend les notations de la question ***16***.

c) Montrer que pour $j \in [\![1, i-1]\!]$, $\mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) = \dfrac{\mathbb{E}_{\overline{A_i}}\left(Y_j \displaystyle\prod_{k \in J_i} (1 - Y_k)\right)}{\mathbb{P}_{\overline{A_i}}(C_i)}$.

Démonstration.
Soit $j \in [\![1, i-1]\!]$.

- Tout d'abord :

$$\begin{aligned}
\mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) &= \frac{\mathbb{P}(\overline{A_i} \cap C_i \cap \overline{A_j})}{\mathbb{P}(\overline{A_i} \cap C_i)} \\
&= \frac{\dfrac{\mathbb{P}(\overline{A_i} \cap C_i \cap \overline{A_j})}{\mathbb{P}(\overline{A_i})}}{\dfrac{\mathbb{P}(\overline{A_i} \cap C_i)}{\mathbb{P}(\overline{A_i})}} \\
&= \frac{\mathbb{P}_{\overline{A_i}}(C_i \cap \overline{A_j})}{\mathbb{P}_{\overline{A_i}}(C_i)}
\end{aligned}$$

- Ensuite :

$$\begin{aligned}
\mathbb{P}_{\overline{A_i}}(C_i \cap \overline{A_j}) &= \mathbb{P}_{\overline{A_i}}\left(\overline{A_j} \cap \left(\bigcap_{k \in J_i} A_k\right)\right) && \textit{(par définition de } C_i\textit{)} \\
&= \mathbb{P}_{\overline{A_i}}\left([Y_j = 1] \cap \left(\bigcap_{k \in J_i} [Y_k = 0]\right)\right) && \textit{(par définition de } \overline{A_j} \textit{ et des } A_k\textit{)}
\end{aligned}$$

- Enfin, remarquons que $Y_j \displaystyle\prod_{k \in J_i} (1 - Y_k)$ est une variable de Bernoulli. À ce titre, et puisque $\mathbb{P}_{\overline{A_i}}$ a les mêmes propriétés que $\mathbb{P}$:

$$\begin{aligned}
\mathbb{E}_{\overline{A_i}}\left(Y_j \prod_{k \in J_i} (1 - Y_k)\right) &= \mathbb{P}_{\overline{A_i}}\left(\left[Y_j \prod_{k \in J_i} (1 - Y_k) = 1\right]\right) \\
&= \mathbb{P}_{\overline{A_i}}\left([Y_j = 1] \cap \left(\bigcap_{k \in J_i} [Y_k = 0]\right)\right)
\end{aligned}$$

$$\boxed{\text{Ainsi, pour tout } j \in [\![1, i-1]\!],\ \mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) = \frac{\mathbb{E}_{\overline{A_i}}\left(Y_j \displaystyle\prod_{k \in J_i} (1 - Y_k)\right)}{\mathbb{P}_{\overline{A_i}}(C_i)}.}$$

□

Commentaire

Détaillons pourquoi $Y_j \prod_{k \in J_i} (1 - Y_k)$ est une variable de Bernoulli.
Cela repose sur deux faits rapides à vérifier :

× si X et Y sont deux variables de Bernoulli, alors XY aussi (car $\{0, 1\}$ est stable par produit),

× si X est une variable de Bernoulli, alors $1 - X$ aussi (car $\{0, 1\}$ est stable par $x \mapsto 1 - x$).

Ainsi, les $(1 - Y_k)$ sont des variables de Bernoulli puis, par produit (fini) de variables de Bernoulli, il suit que $Y_j \prod_{k \in J_i} (1 - Y_k)$ est une variable de Bernoulli.

d) En utilisant l'inégalité de Harris, montrer que pour $j \in [\![1, i-1]\!]$:

$$\mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) \leqslant \mathbb{P}_{\overline{A_i}}(\overline{A_j})$$

Démonstration.

- Commençons par montrer que les variables aléatoires $T_{u,v}$ sont indépendantes pour la probabilité conditionnelle $\mathbb{P}_{\overline{A_i}}$.
  Pour cela, on numérote les variables $T_{u,v}$ de U_1 à U_m où $m = \binom{n}{2}$, de sorte que $Y_i = U_1 U_2 U_3$.
  Il s'agit maintenant de montrer que les variables aléatoires $U_1, \ldots, U_m$ sont indépendantes pour la probabilité conditionnelle $\mathbb{P}_{\overline{A_i}}$.
  Vérifions pour cela que l'on peut appliquer la question ***20.b)*** :
  × les variables aléatoires $U_1, \ldots, U_m$ sont des variables de Bernoulli indépendantes (pour $\mathbb{P}$),
  × $m = \dfrac{n(n-1)}{2} \geqslant 3$ car $n \geqslant 3$,
  × $\overline{A_i} = [Y_i = 1] = [U_1 U_2 U_3 = 1] = \left[\prod_{i \in I} U_i = 1\right]$ où $I = \{1, 2, 3\} \subset [\![1, m]\!]$.
  Les hypothèses sont bien vérifiées donc, d'après la question ***20.b)*** : les variables aléatoires U_k (c'est-à-dire les variables aléatoires $T_{u,v}$) sont indépendantes pour la probabilité conditionnelle $\mathbb{P}_{\overline{A_i}}$.
- Soit $j \in [\![1, i-1]\!]$.
  Rappelons tous les ingrédients nécessaires à l'utilisation de l'inégalité de Harris :
  × les variables $T_{u,v}$ sont finies et indépendantes pour la probabilité conditionnelle $\mathbb{P}_{\overline{A_i}}$ comme vu ci-dessus.
  × d'après la question ***13.a)***, Y_i s'exprime comme une fonction m-croissante sur $\{0, 1\}^m$ des variables aléatoires $T_{u,v}$ pour $u < v$ éléments de $[\![0, n-1]\!]$.
  × toujours d'après la question ***13.a)***, $\prod_{k \in J_i} (1 - Y_k)$ s'exprime comme une fonction m-décroissante des variables aléatoires $T_{u,v}$ pour $u < v$ éléments de $[\![0, n-1]\!]$.
  × l'inégalité de Harris change de sens dans ce contexte, d'après la question ***10.e)***.
  On en déduit alors :

$$\mathbb{E}_{\overline{A_i}}\left(Y_j \prod_{k \in J_i} (1 - Y_k)\right) \leqslant \mathbb{E}_{\overline{A_i}}(Y_j)\, \mathbb{E}_{\overline{A_i}}\left(\prod_{k \in J_i} (1 - Y_k)\right)$$

Or :

- × d'une part :

$$\mathbb{E}_{\overline{A_i}}(Y_j) = \mathbb{P}_{\overline{A_i}}([Y_j = 1]) = \mathbb{P}_{\overline{A_i}}(\overline{A_j})$$

- × d'autre part :

$$\mathbb{E}_{\overline{A_i}}\left(\prod_{k \in J_i}(1 - Y_k)\right) = \mathbb{P}_{\overline{A_i}}\left(\bigcap_{k \in J_i}[Y_k = 0]\right) = \mathbb{P}_{\overline{A_i}}\left(\bigcap_{k \in J_i} A_k\right) = \mathbb{P}_{\overline{A_i}}(C_i)$$

D'où :

$$\mathbb{E}_{\overline{A_i}}\left(Y_j \prod_{k \in J_i}(1 - Y_k)\right) \leqslant \mathbb{P}_{\overline{A_i}}(\overline{A_j})\mathbb{P}_{\overline{A_i}}(C_i)$$

D'après la question ***20.c)***, il reste à diviser par $\mathbb{P}_{\overline{A_i}}(C_i) > 0$:

$$\boxed{\text{On obtient bien : } \mathbb{P}_{\overline{A_i} \cap C_i}(\overline{A_j}) \leqslant \mathbb{P}_{\overline{A_i}}(\overline{A_j}).}$$

□

ESSEC II 2024 - énergie d'une matrice, trace d'une matrice, produit de Kronecker, réduction des matrices symétriques, graphe complet

On s'intéresse dans ce problème à l'énergie d'un graphe qui est définie à partir de l'énergie de sa matrice d'adjacence.

L'énergie d'un graphe a été introduite en 1978 par Ivan Gutman. Ce n'est qu'à partir des années 2000 que des recherches approfondies sur cette notion ont été entreprises.

Aujourd'hui plus de deux articles par semaine sont publiés sur l'énergie des graphes avec de nombreuses applications dans divers domaines scientifiques.

Les parties 2 et 3 sont indépendantes et un bref aide-mémoire Python se trouve en page 6.

Dans tout le problème n est un entier supérieur ou égal à 2.

On rappelle que la notation $\sum\limits_{1\leqslant i,j\leqslant n}$ est analogue à la notation $\sum\limits_{(i,j)\in[\![1,n]\!]^2}$.

Partie 1 - Énergie et trace d'une matrice

1. Soit A une matrice carrée symétrique appartenant à $\mathscr{M}_n(\mathbb{R})$. Justifier l'existence d'une matrice carrée inversible P appartenant à $\mathscr{M}_n(\mathbb{R})$ telle que $P^{-1}AP$ soit une matrice diagonale. Que peut-on dire des éléments diagonaux de $P^{-1}AP$?

Démonstration.
La matrice A est symétrique (réelle) donc est diagonalisable. Ainsi, il existe :

× une matrice $P \in \mathscr{M}_n(\mathbb{R})$ inversible obtenue par concaténation des bases des sous-espaces propres de A,

× une matrice $D \in \mathscr{M}_n(\mathbb{R})$ diagonale dont les coefficients diagonaux sont les valeurs propres de A,

telles que $A = PDP^{-1}$.

$$\boxed{\text{La matrice } P^{-1}AP \text{ est diagonale et ses coefficients diagonaux sont les valeurs propres de } A.}$$

□

► *En notant $\lambda_1, \dots, \lambda_n$ les éléments diagonaux de $P^{-1}AP$, on pose alors $\mathcal{E}(A) = \sum\limits_{k=1}^{n} |\lambda_k|$, on nomme ce réel positif l'énergie de A.*
On admet que cette somme ne dépend pas du choix de P.

2. Montrer que $\mathcal{E}(A) = 3$ si $A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$.

Démonstration.
Tout d'abord, remarquons que A est bien symétrique (réelle) et donc diagonalisable.
De plus, à l'aide de la question ***1***, on peut traduire la définition de l'énergie de $A \in \mathscr{M}_3(\mathbb{R})$ en :

$$\mathcal{E}(A) = \sum_{k=1}^{3} |\lambda_k|$$

où $\lambda_1, \lambda_2, \lambda_3$ sont les valeurs propres de A (éventuellement répétées plusieurs fois selon la dimension du sous-espace propre associé).

Soit $\lambda \in \mathbb{R}$.

$$\begin{aligned}
\operatorname{rg}(A-\lambda I_3) &= \operatorname{rg}\left(\begin{pmatrix} 1-\lambda & -1 & -1 \\ -1 & -\lambda & 0 \\ -1 & 0 & -\lambda \end{pmatrix}\right) \\
&= \operatorname{rg}\left(\begin{pmatrix} -1 & 0 & -\lambda \\ -1 & -\lambda & 0 \\ 1-\lambda & -1 & -1 \end{pmatrix}\right) & L_1 \leftrightarrow L_3 \\
&= \operatorname{rg}\left(\begin{pmatrix} -1 & 0 & -\lambda \\ 0 & -\lambda & \lambda \\ 0 & -1 & -1-\lambda+\lambda^2 \end{pmatrix}\right) & \begin{array}{l} L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 + (1-\lambda)L_1 \end{array} \\
&= \operatorname{rg}\left(\begin{pmatrix} -1 & 0 & -\lambda \\ 0 & -1 & -1-\lambda+\lambda^2 \\ 0 & -\lambda & \lambda \end{pmatrix}\right) & L_2 \leftrightarrow L_3 \\
&= \operatorname{rg}\left(\begin{pmatrix} -1 & 0 & -\lambda \\ 0 & -1 & -1-\lambda+\lambda^2 \\ 0 & 0 & \lambda(\lambda+1)(2-\lambda) \end{pmatrix}\right) & L_3 \leftarrow L_3 - \lambda L_2
\end{aligned}$$

Ainsi :

$$\begin{aligned}
\lambda \in \operatorname{Sp}(A) &\iff A - \lambda I_3 \text{ est non inversible} \\
&\iff \operatorname{rg}(A-\lambda I_3) < 3 \\
&\iff \lambda(\lambda+1)(2-\lambda) = 0
\end{aligned}$$

et donc :

$$\operatorname{Sp}(A) = \{-1, 0, 2\}$$

Puisque $A \in \mathscr{M}_3(\mathbb{R})$ et possède 3 valeurs propres distinctes, chaque sous-espace propre est de dimension 1. En effet, ils sont tous les trois de dimension au moins 1. Si l'un d'entre eux était de dimension au moins 2, on pourrait alors construire par théorème de concaténation une famille libre de cardinal au moins 4 dans $\mathscr{M}_{3,1}(\mathbb{R})$, ce qui serait absurde car $\dim(\mathscr{M}_{3,1}(\mathbb{R})) = 3$.

$$\boxed{\text{On en déduit que } \mathcal{E}(A) = |-1| + |0| + |2| = 1 + 2 = 3.}$$

□

Commentaire

Un œil avisé aura remarqué avant tout calcul que la matrice A n'est pas inversible (car ses deux dernières colonnes sont égales) donc 0 est valeur propre de A.

De plus, la somme des coefficients sur chaque ligne vaut -1 donc $A\begin{pmatrix}1\\1\\1\end{pmatrix} = -\begin{pmatrix}1\\1\\1\end{pmatrix}$ et $\begin{pmatrix}1\\1\\1\end{pmatrix} \neq 0_{\mathscr{M}_{3,1}(\mathbb{R})}$ donc -1 est valeur propre de A.
Avoir ces deux informations en tête permet de détecter une éventuelle erreur de calcul lors du pivot de Gauss effectué afin de déterminer le spectre de A.
Allons même un peu plus loin. À la question ***4.c)***, on démontre que deux matrices semblables ont même trace. Une conséquence de ce résultat est que la trace d'une matrice est la somme de ses valeurs propres (comptées avec multiplicité selon les dimensions des sous-espaces propres). Puisqu'on sait déjà que -1 et 0 sont valeurs propres de A, et puisque $\operatorname{tr}(A) = 1 + 0 + 0 = 1$, la dernière valeur propre ne peut être que 2 (on doit résoudre $-1 + 0 + \lambda = 1$).

3. Ecrire une fonction **Python** `energie(A)` qui renvoie l'énergie de la matrice symétrique représentée par le tableau `numpy A`.

Démonstration.
On utilise la fonction `al.eigvals` proposée en aide-mémoire pour avoir accès aux coefficients d'une matrice diagonale semblable à A, c'est-à-dire aux valeurs propres de A.

```
1 def energie(A)
2     D = al.eigvals(A) # tableau des valeurs propres de A
3     return np.sum(np.abs(D))
```

Détaillons les trois étapes du calcul de l'énergie de A :

× à la ligne 2, on crée une variable `D` contenant le tableau `[`λ_1`,`λ_2`,...,`λ_n`]` où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de A (comptées avec multiplicité selon les dimensions des sous-espaces propres).

× à la ligne 3, `np.abs(D)` contient le tableau `[`$|\lambda_1|$`,`$|\lambda_2|$`,...,`$|\lambda_n|$`]` puis `np.sum(np.abs(D))` renvoie la somme des éléments de ce tableau, c'est à dire l'énergie de A.

□

▸ *Si $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$ est une matrice carrée appartenant à $\mathscr{M}_n(\mathbb{R})$, on définit la trace de A, notée* $\mathrm{tr}(A)$ *par :*

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{i,i}$$

4. Notons $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$ et $B = (b_{i,j})_{1\leqslant i,j\leqslant n}$.

a) Montrer que : $\mathrm{tr}(AB) = \sum\limits_{i=1}^{n} \left(\sum\limits_{j=1}^{n} a_{i,j} b_{j,i} \right) = \sum\limits_{1\leqslant i,j\leqslant n} a_{i,j} b_{j,i}$.

Démonstration.
Par définition du produit matriciel : $AB = (c_{i,j})_{1\leqslant i,j\leqslant n}$ où, pour tout $(i,j) \in [\![1,n]\!]^2$:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Ainsi :

$$\begin{aligned}
\mathrm{tr}(AB) &= \sum_{i=1}^{n} c_{i,i} && \textit{(par définition de la trace)} \\
&= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{i,k} b_{k,i} \right) \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} b_{j,i} \right) \\
&= \sum_{(i,j)\in[\![1,n]\!]^2} a_{i,j} b_{j,i} \\
&= \sum_{1\leqslant i,j\leqslant n} a_{i,j} b_{j,i}
\end{aligned}$$

$$\boxed{\text{On a bien : } \mathrm{tr}(AB) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} b_{j,i} \right) = \sum_{1\leqslant i,j\leqslant n} a_{i,j} b_{j,i}.}$$

□

b) En déduire que : $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ et $\mathrm{tr}({}^tAA) = \sum\limits_{1 \leqslant i,j \leqslant n} a_{j,i}^2$.

Que peut-on dire de A si $\mathrm{tr}({}^tAA) = 0$?

Démonstration.

- Tout d'abord :

$$\begin{aligned}
\mathrm{tr}(AB) &= \sum_{1 \leqslant i,j \leqslant n} a_{i,j} b_{j,i} && \textit{(d'après la question } \textbf{\textit{4.a)}}\textit{)} \\
&= \sum_{1 \leqslant i,j \leqslant n} b_{j,i} a_{i,j} \\
&= \sum_{1 \leqslant j,i \leqslant n} b_{j,i} a_{i,j} \\
&= \mathrm{tr}(BA) && \textit{(d'après la question } \textbf{\textit{4.a)}}\textit{)}
\end{aligned}$$

$$\boxed{\mathrm{tr}(AB) = \mathrm{tr}(BA)}$$

- Ensuite, en notant, pour tout $M \in \mathscr{M}_n(\mathbb{R})$ et pour tout $(i,j) \in [\![1,n]\!]^2$, $[M]_{i,j}$ le coefficient à la position (i,j) d'une matrice M :

$$\begin{aligned}
\mathrm{tr}({}^tAA) &= \sum_{1 \leqslant i,j \leqslant n} [{}^tA]_{i,j} [A]_{j,i} && \textit{(d'après la question } \textbf{\textit{4.a)}}\textit{)} \\
&= \sum_{1 \leqslant i,j \leqslant n} [A]_{j,i} [A]_{j,i} && \textit{(par définition de } {}^tA\textit{)} \\
&= \sum_{1 \leqslant i,j \leqslant n} a_{j,i} a_{j,i} \\
&= \sum_{1 \leqslant i,j \leqslant n} a_{j,i}^2
\end{aligned}$$

$$\boxed{\mathrm{tr}({}^tAA) = \sum_{1 \leqslant i,j \leqslant n} a_{j,i}^2}$$

- Supposons que $\mathrm{tr}({}^tAA) = 0$.
  Alors $\sum\limits_{1 \leqslant i,j \leqslant n} a_{j,i}^2 = 0$.
  Or, une somme de termes positifs est nulle si et seulement si tous les termes de la somme sont nuls.
  On en déduit :
  $$\forall (i,j) \in [\![1,n]\!]^2,\ a_{j,i}^2 = 0$$
  et donc :
  $$\forall (i,j) \in [\![1,n]\!]^2,\ a_{j,i} = 0$$
  $$\boxed{\text{Ainsi : } A = 0_{\mathscr{M}_n(\mathbb{R})}.}$$

$\square$

c) Si A et B sont semblables, montrer que $\text{tr}(A) = \text{tr}(B)$.

Démonstration.
Supposons A et B semblables. Alors il existe une matrice $P \in \mathscr{M}_n(\mathbb{R})$ inversible telle que $A = PBP^{-1}$. On a alors :

$$\begin{aligned} \text{tr}(A) &= \text{tr}\left((PB)P^{-1}\right) \\ &= \text{tr}\left(P^{-1}(PB)\right) && \textit{(d'après la question 4.b))} \\ &= \text{tr}(B) && \textit{(car } P^{-1}P = I_n\textit{)} \end{aligned}$$

$$\boxed{\text{On a bien : } \text{tr}(A) = \text{tr}(B).}$$

□

5. Dans cette question $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ est symétrique et on utilise les notations de la question ***1***.

a) Montrer que $|\text{tr}(A)| \leqslant \mathcal{E}(A)$.

Démonstration.
Avec les notations de la question ***1***, posons : $D = P^{-1}AP$. On a vu que D était une matrice diagonale dont les coefficients diagonaux sont les valeurs propres de A. Notons $\lambda_1, \ldots, \lambda_n$ les coefficients diagonaux de D (c'est-à-dire les valeurs propres de A).
On a alors :

$$\begin{aligned} \text{tr}(A) &= \text{tr}(D) && \textit{(d'après la question 4.c), car A et D sont semblables)} \\ &= \sum_{k=1}^{n} \lambda_k \end{aligned}$$

Or, par inégalité triangulaire :

$$\left| \sum_{k=1}^{n} \lambda_k \right| \leqslant \sum_{k=1}^{n} |\lambda_k|$$

$$\boxed{\text{Cette dernière inégalité se réécrit : } |\text{tr}(A)| \leqslant \mathcal{E}(A).}$$

□

b) Justifier que $\text{tr}(A^2) = \sum\limits_{k=1}^{n} \lambda_k^2$.

Démonstration.
Réutilisons les notations de la question précédente. Tout d'abord :

$$\begin{aligned} A^2 &= (PDP^{-1})^2 \\ &= (PDP^{-1})\ (PDP^{-1}) \\ &= PD(P^{-1}P)DP^{-1} \\ &= PD^2P^{-1} && \textit{(car } P^{-1}P = I_n\textit{)} \end{aligned}$$

et donc les matrices A^2 et D^2 sont semblables. Ainsi, d'après la question ***4.c)***, $\text{tr}(A^2) = \text{tr}(D^2)$.
De plus, la matrice D étant diagonale, la matrice D^2 est encore diagonale et ses coefficients diagonaux sont : $\lambda_1^2, \ldots, \lambda_n^2$.

$$\boxed{\text{D'où : } \text{tr}(A^2) = \text{tr}(D^2) = \sum_{k=1}^{n} \lambda_k^2.}$$

□

6. Dans la console **Python**, on obtient :

```
> > > energie(3*np.eye(3)-np.ones([3,3]))
6.0
```

a) Déterminer quelle est la matrice A associée au tableau `3*np.eye(3)-np.ones([3,3])`.

Démonstration.
Détaillons la construction de la matrice A :

× La commande `np.eye(3)` crée un tableau `numpy` à 3 lignes et 3 colonnes ne contenant que des 1 sur sa diagonale et des 0 ailleurs. Autrement dit, cette commande renvoie la matrice I_3.

× La commande `np.ones([3,3])` crée un tableau `numpy` à 3 lignes et 3 colonnes ne contenant que des 1. Autrement dit, cette commande renvoie la matrice $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

$$\boxed{\text{Donc : } A = 3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}}$$

□

b) En calculant $\mathrm{tr}(A)$, $\mathrm{tr}(A^2)$ et en déterminant une valeur propre évidente de A, expliciter son spectre et retrouver son énergie.

Démonstration.
Tout d'abord : $\mathrm{tr}(A) = 6$. De plus :

$$A^2 = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{pmatrix}$$

$$\boxed{\text{D'où : } \mathrm{tr}(A^2) = 18.}$$

Ensuite, on remarque :

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

et puisque $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \neq 0_{\mathscr{M}_{n,1}(\mathbb{R})}$, on en déduit que 0 est valeur propre de A.

La matrice A étant symétrique, elle est diagonalisable et admet 3 valeurs propres (comptées avec multiplicité). Notons α et β les deux valeurs propres restantes. D'après les questions ***5.a)*** et ***5.b)***, on a :

$$\left\{ \begin{array}{ccccccccc} 0 & + & \alpha & + & \beta & = & \mathrm{tr}(A) & = & 6 \\ 0 & + & \alpha^2 & + & \beta^2 & = & \mathrm{tr}(A^2) & = & 18 \end{array} \right. \quad \text{donc} \quad \left\{ \begin{array}{ccccc} & & \beta & = & 6-\alpha \\ \alpha^2 & + & \beta^2 & = & 18 \end{array} \right.$$

Déterminons α. Il s'agit de résoudre l'équation $\alpha^2 + (6-\alpha)^2 = 18$.
Or :

$$\begin{aligned} \alpha^2 + (6-\alpha)^2 = 18 &\iff \alpha^2 + 36 - 12\alpha + \alpha^2 = 18 \\ &\iff 2\alpha^2 - 12\alpha + 18 = 0 \\ &\iff \alpha^2 - 6\alpha + 9 = 0 \\ &\iff (\alpha - 3)^2 = 0 \\ &\iff \alpha = 3 \end{aligned}$$

On en déduit alors que $\beta = 6 - 3 = 3$.

$$\boxed{\text{On peut conclure : } \mathrm{Sp}(A) = \{0, 3\}.}$$

Puisque la valeur propre 3 apparaît deux fois, on obtient :

$$\mathcal{E}(A) = |0| + |3| + |3| = 6$$

$$\boxed{\text{Notre calcul de l'énergie de } A \text{ est cohérent avec le résultat renvoyé par } \textbf{Python}.}$$

□

Partie 2 - Produit de Kronecker $2 \times n$ de matrices symétriques

- Soit $U = \begin{pmatrix} u & v \\ v & w \end{pmatrix}$ une matrice carrée appartenant à $\mathscr{M}_2(\mathbb{R})$ symétrique et A une matrice carrée appartenant à $\mathscr{M}_n(\mathbb{R})$ symétrique.
  On définit $U * A$ la matrice carrée appartenant à $\mathscr{M}_{2n}(\mathbb{R})$ que l'on peut naturellement représenter ainsi $\begin{pmatrix} uA & vA \\ vA & wA \end{pmatrix}$ (écriture par blocs).

- Par exemple, si $U = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, alors $U * A = \begin{pmatrix} 2A & A \\ A & 0_3 \end{pmatrix}$ par blocs,

  d'où $U * A = \begin{pmatrix} 0 & 2 & 2 & 0 & 1 & 1 \\ 2 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ (0_3 désigne la matrice nulle de $\mathscr{M}_3(\mathbb{R})$).

- Si X est une matrice colonne appartenant à $\mathscr{M}_{2n,1}(\mathbb{R})$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_{2n} \end{pmatrix}$, on écrira $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ avec

  $X_1 = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $X_2 = \begin{pmatrix} x_{n+1} \\ \vdots \\ x_{2n} \end{pmatrix}$.

  On admet qu'alors la matrice colonne $(U * A)X$ de $\mathscr{M}_{2n,1}(\mathbb{R})$ est égale à $\begin{pmatrix} uAX_1 + vAX_2 \\ vAX_1 + wAX_2 \end{pmatrix}$.

7. ***a)*** Écrire une fonction **Python** `prod2K(u,v,w,A)` qui étant donné $U = \begin{pmatrix} u & v \\ v & w \end{pmatrix}$ et A, représentée par un tableau `numpy`, renvoie $U * A$ sous la forme d'un tableau `numpy`.

Démonstration.

```
def prod2K(u,v,w,A):
    n = len(A) # taille de la matrice A
    K = np.zeros([2*n,2*n]) # K sera la matrice U*A
    for i in range(n):
        for j in range(n):
            K[i,j] = u * A[i,j] # partie en haut à gauche
            K[i+n,j] = v * A[i,j] # partie en bas à gauche
            K[i,j+n] = v * A[i,j] # partie en haut à droite
            K[i+n,j+n] = w * A[i,j] # partie en bas à droite
    return K
```

Détaillons la construction de la matrice $U * A = \begin{pmatrix} uA & vA \\ vA & wA \end{pmatrix}$:

× si la matrice A est de taille $n \times n$, alors la matrice $U * A$ est de taille $2n \times 2n$,

× on récupère l'entier n à la ligne 2,

× on crée à la ligne 3 une matrice `K` de taille $2n \times 2n$, initialement remplie de 0,

× on parcourt tous les coefficients de la matrice `K` à l'aide d'une double boucle `for` imbriquée, aux lignes 4 et 5,

× à la ligne 6, on remplit le bloc en haut à gauche (uA) en manipulant les coefficients $K[i,j]$ pour $(i,j) \in [\![0, n-1]\!]^2$ (on rappelle que la numérotation commence toujours à 0 en **Python**),

× à la ligne 7, on remplit le bloc en bas à gauche (vA) en manipulant les coefficients $K[i,j]$ pour $(i,j) \in [\![n, 2n-1]\!] \times [\![0, n-1]\!]$,

× à la ligne 8, on remplit le bloc en haut à droite (vA) en manipulant les coefficients $K[i,j]$ pour $(i,j) \in [\![0, n-1]\!] \times [\![n, 2n-1]\!]$,

× à la ligne 9, on remplit le bloc en bas à droite (wA) en manipulant les coefficients $K[i,j]$ pour $(i,j) \in [\![n, 2n-1]\!]^2$.

Commentaire

On utilise à la ligne 2 la commande `n = len(A)` pour récupérer la taille de la matrice A. Ceci est valide car la matrice A est représentée en **Python** par une liste contenant les lignes de la matrice, elles-même codées par des listes. Ainsi, `n = len(A)` renvoie le nombre de lignes de la matrice A, ce qui correspond bien à sa taille lorsque celle-ci est carrée.
La commande `np.shape(A)` permet également de récupérer la taille de `A` mais cela complique inutilement le code. En effet, il faut écrire dans ce cas `n,n = np.shape(A)`.

□

b) Compléter le code suivant s'affichant dans la console **Python** :

```
> > > prod2K(...,-1,...,...)
array([[-2., 1., 2., -1.],
       [1., -2., -1., 2.],
       [2., -1., -4., 2.],
       [-1., 2., 2., -4.]])
```

Démonstration.
Au vu de la spécification de la fonction, on cherche u, w et A de sorte que la matrice renvoyée soit $U * A$.
D'après la partie du code pré-remplie : $v = -1$.
De plus, en regardant le bloc 2×2 en haut à droite de la matrice $U * A$: $vA = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ et donc $A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$.
Ensuite, en regardant le bloc en haut à gauche de la matrice $U * A$: $uA = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ et donc $u = 1$.
Enfin, en regardant le bloc en bas à droite de la matrice $U * A$: $wA = \begin{pmatrix} -4 & 2 \\ 2 & -4 \end{pmatrix}$ et donc $w = 2$.

Il faut écrire : `prod2K(1,-1,2,np.array([[-2,1],[1,-2]]))`.

□

8. Soit $\begin{pmatrix} a \\ b \end{pmatrix}$ un vecteur propre de U pour la valeur propre λ et X un vecteur propre de A pour μ. On pose $Y = \begin{pmatrix} aX \\ bX \end{pmatrix}$.

a) Établir l'égalité : $(U * A)Y = \mu \begin{pmatrix} (ua+vb)X \\ (va+wb)X \end{pmatrix}$.

Démonstration.
D'après la formule admise dans l'énoncé :

$$\begin{aligned}
(U * A)Y &= \begin{pmatrix} uA(aX) + vA(bX) \\ vA(aX) + wA(bX) \end{pmatrix} \\
&= \begin{pmatrix} ua\,(AX) + vb\,(AX) \\ va\,(AX) + wb\,(AX) \end{pmatrix} \\
&= \begin{pmatrix} ua\,(\mu X) + vb\,(\mu X) \\ va\,(\mu X) + wb\,(\mu X) \end{pmatrix} && \textit{(car X est un vecteur propre de A associé à la valeur propre μ.)} \\
&= \begin{pmatrix} \mu(ua+vb)X \\ \mu(va+wb)X \end{pmatrix} \\
&= \mu \begin{pmatrix} (ua+vb)X \\ (va+wb)X \end{pmatrix}
\end{aligned}$$

$$\boxed{\text{On a bien : } (U * A)Y = \mu \begin{pmatrix} (ua+vb)X \\ (va+wb)X \end{pmatrix}.}$$

□

b) Montrer que Y est un vecteur propre de $U * A$ et préciser pour quelle valeur propre.

Démonstration.

- Commençons par montrer que Y n'est pas le vecteur nul :
  - × le vecteur X n'est pas nul (sinon ce ne serait pas un vecteur propre de A) et donc possède au moins une coordonnée non nulle,
  - × le vecteur $\begin{pmatrix} a \\ b \end{pmatrix}$ n'est pas nul non plus (sinon ce ne serait pas un vecteur propre de U) et donc $a \neq 0$ ou $b \neq 0$.

  Dans tous les cas, Y possède au moins une coordonnée non nulle et donc n'est pas le vecteur nul.
- D'après la question ***8.a)*** :
$$(U * A)Y = \mu \begin{pmatrix} (ua+vb)X \\ (va+wb)X \end{pmatrix}$$
  On doit donc calculer $ua + vb$ et $va + wb$.
- Pour faire ce calcul, on utilise le fait que $\begin{pmatrix} a \\ b \end{pmatrix}$ est un vecteur propre de U pour la valeur propre λ. Plus précisément :
  - × d'une part :
$$U \begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix}$$
  - × d'autre part :
$$U \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ua+vb \\ va+wb \end{pmatrix}$$

On identifie et on obtient :

$$\begin{aligned} ua + vb &= \lambda a \\ va + wb &= \lambda b \end{aligned}$$

- Reprenons la formule de la question ***8.a)*** :

$$\begin{aligned} (U * A)Y &= \mu \begin{pmatrix} (ua+vb)X \\ (va+wb)X \end{pmatrix} \\ &= \mu \begin{pmatrix} \lambda aX \\ \lambda bX \end{pmatrix} \\ &= \mu\lambda \begin{pmatrix} aX \\ bX \end{pmatrix} \\ &= \mu\lambda Y \end{aligned}$$

$$\boxed{\text{On en déduit que } Y \text{ est un vecteur propre de } U * A \text{ pour la valeur propre } \mu\lambda.}$$

□

9. a) Justifier l'existence d'une base $\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right)$ de $\mathscr{M}_{2,1}(\mathbb{R})$ et de $(X_1, \dots, X_n)$ une base de $\mathscr{M}_{n,1}(\mathbb{R})$, formée de vecteurs propres de U et A respectivement.

Démonstration.
La matrice U est symétrique (réelle) donc diagonalisable. Comme $U \in \mathscr{M}_2(\mathbb{R})$, alors :

$$\boxed{\text{il existe une base } \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) \text{ de } \mathscr{M}_{2,1}(\mathbb{R}) \text{ formée de vecteurs propres de } U.}$$

De manière analogue, la matrice A est symétrique (réelle) donc diagonalisable. Comme $A \in \mathscr{M}_n(\mathbb{R})$, alors

$$\boxed{\text{il existe une base } (X_1, \dots, X_n) \text{ de } \mathscr{M}_{n,1}(\mathbb{R}) \text{ formée de vecteurs propres de } A.}$$

□

► *On note γ_1 et γ_2 les valeurs propres associées à $\begin{pmatrix} a \\ b \end{pmatrix}$ et $\begin{pmatrix} c \\ d \end{pmatrix}$ respectivement et $\mu_1, \dots, \mu_n$ celles associées à $X_1, \dots, X_n$ respectivement.*
On pose, pour tout $i \in [\![1, n]\!]$, $Y_i = \begin{pmatrix} aX_i \\ bX_i \end{pmatrix}$ et $Z_i = \begin{pmatrix} cX_i \\ dX_i \end{pmatrix}$.

b) Montrer que la famille $(Y_1, Z_1, \dots, Y_n, Z_n)$ est libre.

Démonstration.
Soit $(y_1, z_1, \dots, y_n, z_n) \in \mathbb{R}^{2n}$. Supposons :

$$y_1 Y_1 + z_1 Z_1 + \cdots + y_n Y_n + z_n Z_n = 0_{\mathscr{M}_{2n,1}(\mathbb{R})} \quad (*)$$

c'est-à-dire :

$$\sum_{i=1}^{n} y_i Y_i + \sum_{i=1}^{n} z_i Z_i = 0_{\mathscr{M}_{2n,1}(\mathbb{R})}$$

Montrons alors : $y_1 = z_1 = \cdots = y_n = z_n = 0$.

Pour cela, écrivons différemment la somme $\sum\limits_{i=1}^{n} y_i Y_i + \sum\limits_{i=1}^{n} z_i Z_i$:

$$\begin{aligned}
\sum_{i=1}^{n} y_i Y_i + \sum_{i=1}^{n} z_i Z_i &= \sum_{i=1}^{n} y_i \begin{pmatrix} aX_i \\ bX_i \end{pmatrix} + \sum_{i=1}^{n} z_i \begin{pmatrix} cX_i \\ dX_i \end{pmatrix} && \textit{(par définition de } Y_i \textit{ et } Z_i\textit{)} \\
&= \sum_{i=1}^{n} \begin{pmatrix} ay_iX_i \\ by_iX_i \end{pmatrix} + \sum_{i=1}^{n} \begin{pmatrix} cz_iX_i \\ dz_iX_i \end{pmatrix} \\
&= \sum_{i=1}^{n} \left(\begin{pmatrix} ay_iX_i \\ by_iX_i \end{pmatrix} + \begin{pmatrix} cz_iX_i \\ dz_iX_i \end{pmatrix} \right) && \textit{(par linéarité de la somme)} \\
&= \sum_{i=1}^{n} \begin{pmatrix} (ay_i + cz_i)X_i \\ (by_i + dz_i)X_i \end{pmatrix} \\
&= \begin{pmatrix} \sum\limits_{i=1}^{n} (ay_i + cz_i)X_i \\ \sum\limits_{i=1}^{n} (by_i + dz_i)X_i \end{pmatrix}
\end{aligned}$$

D'après $(*)$, on en déduit :

$$\sum_{i=1}^{n} (ay_i + cz_i)X_i = 0_{\mathscr{M}_{n,1}(\mathbb{R})} \qquad \text{et} \qquad \sum_{i=1}^{n} (by_i + dz_i)X_i = 0_{\mathscr{M}_{n,1}(\mathbb{R})}$$

Par liberté de la famille $(X_1, \ldots, X_n)$, on peut conclure :

$$\forall i \in [\![1, n]\!], ay_i + cz_i = 0 \qquad \text{et} \qquad \forall i \in [\![1, n]\!], by_i + dz_i = 0$$

Soit $i \in [\![1, n]\!]$. On vient d'obtenir n couples d'égalités dont le i^{e} se traduit de la manière suivante :

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} y_i \\ z_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Or, la famille $\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right)$ est une base donc est libre.

On en déduit que $\text{rg}\left(\begin{pmatrix} a & c \\ b & d \end{pmatrix} \right) = 2$ et donc la matrice $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ est inversible. D'où : $y_i = z_i = 0$.

Ceci étant valable pour tout $i \in [\![1, n]\!]$, on a : $y_1 = z_1 = \cdots = y_n = z_n = 0$.

$$\boxed{\text{La famille } (Y_1, Z_1, \ldots, Y_n, Z_n) \text{ est libre.}}$$

□

c) En déduire que $U * A$ est semblable à la matrice diagonale dont les éléments diagonaux sont $\gamma_1\mu_1, \gamma_2\mu_1, \ldots, \gamma_1\mu_n, \gamma_2\mu_n$.

Démonstration.
Montrons qu'il existe une base de $\mathscr{M}_{2n,1}(\mathbb{R})$ constituée de vecteurs propres de $U * A$ associés respectivement aux valeurs propres $\gamma_1\mu_1, \gamma_2\mu_1, \ldots, \gamma_1\mu_n, \gamma_2\mu_n$.

- Tout d'abord :
  - × la famille $\mathcal{F} = (Y_1, Z_1, \ldots, Y_n, Z_n)$ est libre (d'après la question ***9.b)***),
  - × $\text{Card}(\mathcal{F}) = n + n = 2n = \dim(\mathscr{M}_{2n,1}(\mathbb{R}))$.

  On en déduit que la famille $(Y_1, Z_1, \ldots, Y_n, Z_n)$ est une base de $\mathscr{M}_{2n,1}(\mathbb{R})$.
- Soit $i \in [\![1, n]\!]$. D'après la question ***8.b)***, Y_i est un vecteur propre de $U * A$ associé à la valeur propre $\gamma_1\mu_i$ et Z_i est un vecteur propre de $U * A$ associé à la valeur propre $\gamma_2\mu_i$.

On note P la matrice dont les colonnes sont, dans cet ordre, $Y_1, Z_1, \ldots, Y_n, Z_n$. La matrice P est inversible car la famille $(Y_1, Z_1, \ldots, Y_n, Z_n)$ est libre. D'après la formule de changement de base, la matrice $D = P^{-1}(U * A)P$ est diagonale et ses coefficients diagonaux sont, dans cet ordre, $\gamma_1\mu_1, \gamma_2\mu_1, \ldots, \gamma_1\mu_n, \gamma_2\mu_n$.

$$\boxed{\text{$U * A$ est semblable à la matrice diagonale D dont les éléments diagonaux sont $\gamma_1\mu_1, \gamma_2\mu_1, \ldots, \gamma_1\mu_n, \gamma_2\mu_n$.}}$$ □

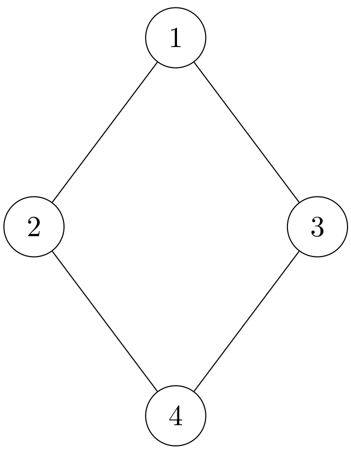
d) En conclure que $\mathcal{E}(U * A) = \mathcal{E}(U) \times \mathcal{E}(A)$.

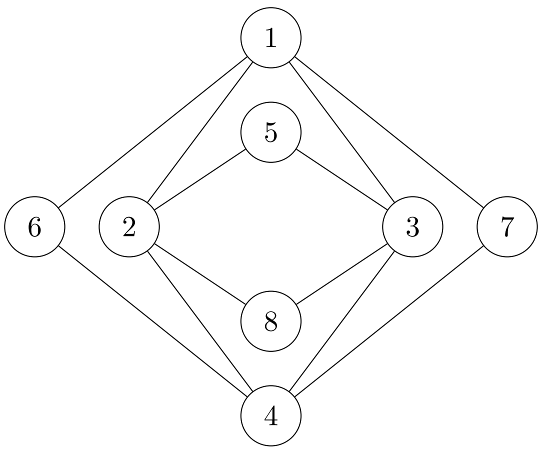
Démonstration.
D'après la question ***9.c)***, $U*A$ est semblable à la matrice diagonale D dont les éléments diagonaux sont $\gamma_1\mu_1, \gamma_2\mu_1, \ldots, \gamma_1\mu_n, \gamma_2\mu_n$. Ainsi :

$$\begin{aligned}
\mathcal{E}(U * A) &= \sum_{i=1}^{n} |\gamma_1\mu_i| + \sum_{i=1}^{n} |\gamma_2\mu_i| && \textit{(par définition de l'énergie d'une matrice)} \\
&= \sum_{i=1}^{n} (|\gamma_1|\,|\mu_i| + |\gamma_2|\,|\mu_i|) \\
&= \sum_{i=1}^{n} (|\gamma_1| + |\gamma_2|)\,|\mu_i| \\
&= (|\gamma_1| + |\gamma_2|) \sum_{i=1}^{n} |\mu_i| \\
&= \mathcal{E}(U) \times \sum_{i=1}^{n} |\mu_i| && \textit{(car la matrice U possède 2 valeurs propres comptées avec multiplicité, qui sont γ_1 et γ_2)} \\
&= \mathcal{E}(U) \times \mathcal{E}(A) && \textit{(car la matrice A possède n valeurs propres comptées avec multiplicité, qui sont $\mu_1, \ldots, \mu_n$)}
\end{aligned}$$

$$\boxed{\text{On en déduit : } \mathcal{E}(U * A) = \mathcal{E}(U) \times \mathcal{E}(A).}$$ □

10. *Un exemple* - On considère un graphe G_n, non orienté et sans boucle, dont les sommets sont $1, \ldots, n$ et l'ensemble des arêtes est noté $\mathcal{A}_n$. On définit le graphe G_{2n} dont les sommets sont $1, \ldots, 2n$ et les arêtes sont celles de G_n, ainsi que, pour toute arête $\{i, j\} \in \mathcal{A}_n$, les arêtes $\{i + n, j\}$ et $\{i, j + n\}$. Voici un exemple de représentation de G_4 et G_8 :





On note A_n et A_{2n} les matrices d'adjacence de G_n et G_{2n}.

a) Déterminer U telle que $A_{2n} = U * A_n$.

Démonstration.

- Commençons par écrire les matrices A_4 et A_8 obtenues par lecture des graphes donnés en exemple :

$$A_4 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \qquad \text{et} \qquad A_8 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

En regardant la matrice A_8 par blocs, on remarque que $A_8 = \begin{pmatrix} A_4 & A_4 \\ A_4 & 0_4 \end{pmatrix}$, c'est-à-dire :

$$A_8 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} * A_4$$

- Plus généralement, montrons que $A_{2n} = \begin{pmatrix} A_n & A_n \\ A_n & 0_n \end{pmatrix}$ (écriture par blocs), quelque soit le graphe G_n. Notons $A_{2n} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$ où, pour tout $i \in [\![1, 4]\!]$, $B_i \in \mathscr{M}_n(\mathbb{R})$.

  × Pour toute arête $\{i, j\} \in \mathcal{A}_n$, on a $i + n \geqslant n + 1$ et $j + n \geqslant n + 1$ donc les nouvelles arêtes ajoutées dans G_{2n} ne relient jamais deux sommets parmi les sommets $1, \dots, n$. Ceci prouve que $B_1 = A_n$ (on a exactement les mêmes arêtes dans G_n et dans G_{2n} entre les sommets $1, \dots, n$).

  × Par construction, pour tout $(i, j) \in [\![1, n]\!]^2$, $[A_{2n}]_{i,j+n} = [A_n]_{i,j}$.
  En effet, si $\{i, j\}$ est une arête de G_n (*i.e.* si $[A_n]_{i,j} = 1$), alors $\{i, j + n\}$ est une arête de G_{2n} (*i.e.* $[A_{2n}]_{i,j+n} = 1$).
  Réciproquement, si $\{i, j\}$ n'est pas une arête de G_n (*i.e.* si $[A_n]_{i,j} = 0$), alors $\{i, j+n\}$ n'est pas une arête de G_{2n} (*i.e.* $[A_{2n}]_{i,j+n} = 0$).
  On a donc $B_2 = A_n$.

  × La matrice A_{2n} étant une matrice d'adjacence d'un graphe non orienté, elle est symétrique. Ainsi : $B_3 = B_2 = A_n$.

  × Par construction, il n'y a aucune arête dans G_{2n} reliant deux sommets x et y vérifiant $x \geqslant n + 1$ et $y \geqslant n + 1$. Ainsi : $B_4 = 0_n$.

$$\boxed{\text{On a démontré que } A_{2n} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} * A_n. \text{ On pose donc } U = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.}$$

□

b) En déduire que $\mathcal{E}(A_{2n}) = \sqrt{5}\ \mathcal{E}(A_n)$.

Démonstration.

- Tout d'abord, en posant $U = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$:

$$\begin{aligned} \mathcal{E}(A_{2n}) &= \mathcal{E}(U * A_n) && \textit{(d'après la question } \textbf{\textit{10.a)}}\textit{)} \\ &= \mathcal{E}(U) \times \mathcal{E}(A_n) && \textit{(d'après la question } \textbf{\textit{9.d)}}\textit{)} \end{aligned}$$

- Calculons maintenant $\mathcal{E}(U)$.
  Soit $\lambda \in \mathbb{R}$.

$$\begin{aligned}\det(U - \lambda I_2) &= \begin{vmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{vmatrix} \\ &= -\lambda(1-\lambda) - 1 \\ &= \lambda^2 - \lambda - 1\end{aligned}$$

  Le discriminant de ce trinôme du second degré est $\Delta = 1 + 4 = 5$.
  Ses deux racines sont $\lambda_1 = \dfrac{1+\sqrt{5}}{2} > 0$ et $\lambda_2 = \dfrac{1-\sqrt{5}}{2} < 0$.
  D'où : $\mathrm{Sp}(U) = \left\{\dfrac{1+\sqrt{5}}{2}, \dfrac{1-\sqrt{5}}{2}\right\}$.

On en déduit :

$$\mathcal{E}(U) = |\lambda_1| + |\lambda_2| = \frac{1+\sqrt{5}}{2} + \frac{\sqrt{5}-1}{2} = \sqrt{5}$$

$$\boxed{\text{D'où : } \mathcal{E}(A_{2n}) = \sqrt{5}\ \mathcal{E}(A_n).}$$

□

Partie 3 - Encadrement de l'énergie d'une matrice d'adjacence

Soit m, n et p des entiers tels que $m \geqslant 1$, $n \geqslant p \geqslant 2$.

On suppose dans cette partie que $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ est la matrice d'adjacence d'un graphe $G(A)$, non orienté sans boucle, à n sommets $1, \dots, n$, m arêtes et $n - p$ sommets isolés, c'est-à-dire de degré 0.

On note $(E_1, \dots, E_n)$ la base canonique de $\mathscr{M}_{n,1}(\mathbb{R})$ et I l'ensemble des sommets non isolés de $G(A)$.

11. a) Montrer que $\sum\limits_{1 \leqslant i,j \leqslant n} a_{i,j} = \sum\limits_{(i,j) \in I^2} a_{i,j} = 2m$.

Démonstration.
Soit $i_0 \in [\![1, n]\!]$. Si i_0 est un sommet isolé (c'est-à-dire si $i_0 \notin I$), alors, pour tout $j \in [\![1, n]\!]$, $a_{i_0,j} = 0$ et $a_{j,i_0} = 0$ (i_0 n'a aucun voisin et le graphe $G(A)$ est non orienté). Ainsi :

$$\begin{aligned}\sum_{1 \leqslant i,j \leqslant n} a_{i,j} &= \sum_{(i,j) \in [\![1,n]\!]^2} a_{i,j} \\ &= \sum_{\substack{(i,j) \in [\![1,n]\!]^2 \\ i \text{ et } j \text{ non isolés}}} a_{i,j} + \sum_{\substack{(i,j) \in [\![1,n]\!]^2 \\ i \text{ isolé ou } j \text{ isolé}}} a_{i,j} \\ &= \sum_{(i,j) \in I^2} a_{i,j} + \sum_{i \notin I \texttt{ OU } j \notin I} a_{i,j} \\ &= \sum_{(i,j) \in I^2} a_{i,j} + \sum_{i \notin I \texttt{ OU } j \notin I} 0 \qquad \textit{(d'après la remarque précédente)} \\ &= \sum_{(i,j) \in I^2} a_{i,j}\end{aligned}$$

$$\boxed{\text{D'où : } \sum_{1 \leqslant i,j \leqslant n} a_{i,j} = \sum_{(i,j) \in I^2} a_{i,j}.}$$

De plus :

$$\begin{aligned}
\sum_{(i,j)\in I^2} a_{i,j} &= \sum_{\substack{(i,j)\in I^2 \\ i=j}} a_{i,j} + \sum_{\substack{(i,j)\in I^2 \\ i\neq j}} a_{i,j} \\
&= \sum_{i\in I} a_{i,i} + \sum_{\substack{(i,j)\in I^2 \\ i<j}} a_{i,j} + \sum_{\substack{(i,j)\in I^2 \\ j<i}} a_{i,j} \\
&= \sum_{i\in I} a_{i,i} + \sum_{\substack{(i,j)\in I^2 \\ i<j}} a_{i,j} + \sum_{\substack{(i,j)\in I^2 \\ j<i}} a_{j,i} && \textit{(car } a_{i,j} = a_{j,i}\textit{)} \\
&= \sum_{i\in I} a_{i,i} + \sum_{\substack{(i,j)\in I^2 \\ i<j}} a_{i,j} + \sum_{\substack{(i,j)\in I^2 \\ i<j}} a_{i,j} && \textit{(car les variables } i \textit{ et } j \textit{ sont muettes)} \\
&= \sum_{i\in I} a_{i,i} + 2 \sum_{\substack{(i,j)\in I^2 \\ i<j}} a_{i,j}
\end{aligned}$$

- × Le graphe $G(A)$ est sans boucle, donc : $\forall i \in I, a_{i,i} = 0$. D'où :

$$\sum_{i\in I} a_{i,i} = 0$$

- × Le graphe $G(A)$ possède m arêtes et la somme $\sum\limits_{\substack{(i,j)\in I^2 \\ i<j}} a_{i,j}$ compte précisément le nombre d'arêtes (la restriction $i < j$ permet de ne pas compter deux fois la même arête). D'où :

$$\sum_{\substack{(i,j)\in I^2 \\ i<j}} a_{i,j} = m$$

$$\boxed{\text{On a bien : } \sum_{(i,j)\in I^2} a_{i,j} = 2m.}$$

□

b) Établir que : $1 \leqslant \dfrac{2m}{p} \leqslant p - 1$.

Démonstration.
D'après la formule d'Euler (dite des poignées de mains) :

$$\sum_{i=1}^{n} \deg(i) = 2m$$

Remarquons que $\deg(i) = 0$ si et seulement si $i \notin I$. On en déduit :

$$\sum_{i=1}^{n} \deg(i) = \sum_{\substack{i=1 \\ i\in I}}^{n} \deg(i) + \sum_{\substack{i=1 \\ i\notin I}}^{n} \deg(i) = \sum_{i\in I} \deg(i)$$

Puisqu'il y a $n - p$ sommets isolés, il y a donc p sommets non isolés. D'où : $\mathrm{Card}(I) = p$.
Considérons $i \in I$ un sommet non isolé.

- × Tout d'abord, i possède au moins un voisin. D'où : $\deg(i) \geqslant 1$ $(*)$.
- × Ensuite, i peut avoir au plus $p-1$ voisins (p voisins potentiels que sont les sommets non isolés, auxquels on doit retirer le sommet i puisque le graphe $G(A)$ est sans boucle).
  D'où : $\deg(i) \leqslant p - 1$ $(**)$.

On en déduit :

$$\begin{aligned} p &= \operatorname{Card}(I) \\ &= \sum_{i\in I} 1 \\ &\leqslant \sum_{i\in I} \deg(i) && \textit{(d'après } (*)\textit{)} \\ &\leqslant \sum_{i\in I} (p-1) && \textit{(d'après } (**)\textit{)} \\ &= p(p-1) \end{aligned}$$

D'où :

$$p \leqslant 2m \leqslant p(p-1)$$

et puisque $p \geqslant 2 > 0$, on a bien :

$$\boxed{1 \leqslant \frac{2m}{p} \leqslant p-1}$$

□

12. a) Justifier qu'il existe une matrice carrée inversible P appartenant à $\mathscr{M}_n(\mathbb{R})$ telle que $P^{-1}AP$ soit une matrice diagonale différente de la matrice nulle.

Démonstration.
La matrice A est symétrique (réelle) donc elle est diagonalisable. A ce titre, il existe une matrice carrée inversible P appartenant à $\mathscr{M}_n(\mathbb{R})$ telle que $D = P^{-1}AP$ soit une matrice diagonale.
Si cette matrice diagonale D était nulle, alors, comme $A = PDP^{-1}$, A serait nulle également.
Or, le graphe $G(A)$ possède au moins un sommet non isolé (puisque $p \geqslant 2$) et donc sa matrice d'adjacence n'est pas nulle.

$$\boxed{\text{Il existe une matrice carrée inversible } P \text{ appartenant à } \mathscr{M}_n(\mathbb{R}) \text{ telle que } P^{-1}AP \text{ soit une matrice diagonale différente de la matrice nulle.}}$$

□

▶ Dans la suite on note D cette matrice diagonale.

b) En déduire que $\operatorname{tr}(D) = 0$, $\operatorname{tr}(D^2) = \sum\limits_{1\leqslant i,j\leqslant n} a_{i,j}^2 = 2m$.

Démonstration.

- D'après la question ***4.c)*** :

$$\begin{aligned} \operatorname{tr}(D) &= \operatorname{tr}(A) && \textit{(car } A \textit{ et } D \textit{ sont semblables)} \\ &= \sum_{i=1}^{n} a_{i,i} \\ &= \sum_{i=1}^{n} 0 && \textit{(car le graphe } G(A) \textit{ est sans boucle, donc, pour tout } i \in [\![1,n]\!]\textit{, le sommet } i \textit{ n'est pas voisin de lui-même et donc } a_{i,i}=0\textit{)} \\ &= 0 \end{aligned}$$

$$\boxed{\text{D'où : } \operatorname{tr}(D) = 0.}$$

- D'après la question ***4.a)*** appliquée avec $B = A$:

$$\begin{aligned}
\operatorname{tr}(A^2) &= \sum_{1 \leqslant i,j \leqslant n} a_{i,j} a_{j,i} \\
&= \sum_{1 \leqslant i,j \leqslant n} a_{i,j}^2 && \textit{(}G(A) \textit{ est non orienté donc } a_{j,i} = a_{i,j}\textit{)} \\
&= \sum_{1 \leqslant i,j \leqslant n} a_{i,j} && \textit{(}a_{i,j} \in \{0,1\} \textit{ donc } a_{i,j}^2 = a_{i,j}\textit{)} \\
&= 2m && \textit{(d'après la question } \boldsymbol{11.a)}\textit{)}
\end{aligned}$$

De plus, D^2 et A^2 sont semblables.

$$\boxed{\text{D'après la question } \textbf{\textit{4.c)}} : \operatorname{tr}(D^2) = \operatorname{tr}(A^2) = 2m.}$$

□

▶ *On suppose dans la suite que P est telle que les éléments diagonaux de D, $\lambda_1, \ldots, \lambda_n$ vérifient $|\lambda_1| \geqslant \cdots \geqslant |\lambda_n|$. On pose $\theta = \max\limits_{1 \leqslant k \leqslant n} \lambda_k$.*

Commentaire

Montrons qu'un tel choix de matrice P est toujours possible. Partons d'une matrice P inversible telle que $D = P^{-1}AP$ est diagonale et notons $\lambda_1, \ldots, \lambda_n$ les éléments diagonaux de D.
Il existe σ une bijection de $[\![1,n]\!]$ dans $[\![1,n]\!]$ telle que :

$$\left|\lambda_{\sigma(1)}\right| \geqslant \cdots \geqslant \left|\lambda_{\sigma(n)}\right|$$

Numérotons $C_1, \ldots, C_n$ les colonnes de la matrice P et notons $\tilde{P}$ la matrice obtenue en concaténant les colonnes $C_{\sigma(1)}, \ldots, C_{\sigma(n)}$ dans cet ordre. La matrice $\tilde{P}$ est, tout comme la matrice P, inversible : elle a été obtenue en permutant les colonnes de P.
Par formule de changement de base, la matrice $\tilde{D} = \tilde{P}^{-1}A\tilde{P}$ est diagonale et ses coefficients diagonaux sont, dans cet ordre, $\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(n)}$. Ainsi, la matrice $\tilde{P}$ convient.

13. a) Soit k un sommet isolé. Montrer que $E_k \in \ker(A)$.
En déduire que $\dim(\ker(A)) \geqslant n - p$ puis que, si $p < n$, $\lambda_{p+1} = \cdots = \lambda_n = 0$.

Démonstration.

- Le vecteur E_k étant le k^{e} vecteur de la base canonique, il suit que AE_k est la k^{e} colonne de la matrice A. Ainsi :

$$AE_k = \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{n,k} \end{pmatrix}$$

Or, puisque k est un sommet isolé, pour tout $j \in [\![1,n]\!]$, $a_{j,k} = 0$.

$$\boxed{\text{On a } AE_k = 0_{\mathscr{M}_{n,1}(\mathbb{R})} \text{ donc } E_k \in \ker(A).}$$

- Notons $\overline{I}$ le complémentaire de I dans $[\![1,n]\!]$. Autrement dit, $\overline{I}$ est l'ensemble des sommets isolés de $G(A)$.
Les propriétés suivantes sont vérifiées :
  × la famille $\mathcal{F} = (E_k)_{k \in \overline{I}}$ est libre comme sous-famille de la base canonique de $\mathscr{M}_{n,1}(\mathbb{R})$ qui est libre,

× $\text{Card}(\mathcal{F}) = \text{Card}\left(\overline{I}\right) = n - p$ (il y a $n-p$ sommets isolés),

× d'après le point précédent, pour tout $k \in \overline{I}$, $E_k \in \ker(A)$.

On a donc construit une famille libre de cardinal $n-p$ dans $\ker(A)$, ce qui implique :

$$\boxed{\dim(\ker(A)) \geqslant n - p}$$

- Supposons : $p < n$. Alors $\dim(\ker(A)) > 0$ et donc 0 est valeur propre de A.
  Plus précisément, $\dim(\ker(A)) \geqslant n-p$ donc 0 apparaît au moins $n-p$ fois sur la diagonale de la matrice D.
  Ainsi, parmi les nombres $\lambda_1, \dots, \lambda_n$, au moins $n-p$ d'entre eux sont nuls.
  Notons q le premier indice tel que $\lambda_q = 0$. Autrement dit : $q = \min\left(\{k \in [\![1,n]\!] \mid \lambda_k = 0\}\right)$.

  × Par construction de P, on a $|\lambda_1| \geqslant \cdots \geqslant |\lambda_n|$. En particulier :

  $$0 = |\lambda_q| \geqslant |\lambda_{q+1}| \geqslant \cdots \geqslant |\lambda_n| \geqslant 0$$

  d'où

  $$\lambda_q = \lambda_{q+1} = \cdots = \lambda_n = 0$$

  × Supposons $q > p+1$. Alors, par minimalité de q, on a $\lambda_{q-1} \neq 0$ (on a bien $q-1 \in [\![1,n]\!]$ puisque $n \geqslant q > p+1 \geqslant 3$). On en déduit que

  $$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_{q-1}| > 0$$

  et donc

  $$\forall k \in [\![1, q-1]\!],\ \lambda_k \neq 0$$

  On a donc au plus $n-q+1$ termes nuls sur la diagonale de D (les nombres $\lambda_q, \dots, \lambda_n$).
  Or, $q > p+1$ donc $n-q+1 < n-p$. Cela contredit le fait que 0 apparaît au moins $n-p$ fois sur la diagonale de D. C'est donc absurde.

  $$\boxed{\text{D'où } q \leqslant p+1 \text{ et } \lambda_{p+1} = \cdots = \lambda_n = 0.}$$

□

Commentaire

Nous avons utilisé dans la preuve précédente le fait suivant : le vecteur E_k étant le k^{e} vecteur de la base canonique, il suit que AE_k est la k^{e} colonne de la matrice A. Il s'agit d'un raccourci calculatoire très pratique à connaître. Détaillons une preuve de ce fait.
Soit $k \in [\![1,n]\!]$. Par produit matriciel :

$$\begin{aligned}
AE_k &= \begin{pmatrix} \sum\limits_{j=1}^{n} [A]_{1,j}[E_k]_j \\ \vdots \\ \sum\limits_{j=1}^{n} [A]_{n,j}[E_k]_j \end{pmatrix} \\
&= \begin{pmatrix} [A]_{1,k}[E_k]_k \\ \vdots \\ [A]_{n,k}[E_k]_k \end{pmatrix} && \textit{(car, pour tout } j \in [\![1,n]\!] \setminus \{k\},\ [E_k]_j = 0\textit{)} \\
&= \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{n,k} \end{pmatrix} && \textit{(car, pour tout } i \in [\![1,n]\!],\ [A]_{i,k} = a_{i,k} \textit{ et } [E_k]_k = 1\textit{)}
\end{aligned}$$

b) Montrer que $\sum\limits_{k=1}^{p} \lambda_k^2 = 2m$ puis que $0 < \theta < \sqrt{2m}$.

Démonstration.

- Tout d'abord :

$$\begin{aligned} 2m &= \mathrm{tr}(D^2) && \textit{(d'après la question } \textbf{\textit{12.b)}}\textit{)} \\ &= \sum_{k=1}^{n} \lambda_k^2 \\ &= \sum_{k=1}^{p} \lambda_k^2 + \sum_{k=p+1}^{n} \lambda_k^2 \\ &= \sum_{k=1}^{p} \lambda_k^2 && \textit{(d'après la question } \textbf{\textit{13.a)}}\textit{)} \end{aligned}$$

$$\boxed{\text{D'où : } \sum_{k=1}^{p} \lambda_k^2 = 2m.}$$

- Montrons : $\theta > 0$. Pour cela, raisonnons par l'absurde et supposons : $\theta \leqslant 0$.
  Par définition du maximum, on a : pour tout $k \in [\![1, n]\!]$, $\lambda_k \leqslant \theta \leqslant 0$.
  Par sommation de ces inégalités, on obtient :

$$\mathrm{tr}(D) = \sum_{k=1}^{n} \lambda_k \leqslant 0$$

  Or, d'après la question ***12.b)***, $\mathrm{tr}(D) = 0$. Il y a donc égalité dans l'inégalité précédente, ce qui ne peut arriver que si tous les λ_k sont nuls. Ceci implique que D est la matrice nulle, ce qui contredit la question ***12.a)***. C'est absurde.

$$\boxed{\text{D'où : } \theta > 0.}$$

- Notons q un entier de $[\![1, n]\!]$ vérifiant $\lambda_q = \theta$.
  - × Montrons qu'il existe $k \in [\![1, n]\!] \setminus \{q\}$ tel que $\lambda_k \neq 0$ en raisonnant par l'absurde.
    Supposons que pour tout $k \in [\![1, n]\!] \setminus \{q\}$, $\lambda_k = 0$.
    On a alors, d'après le point précédent :

$$0 = \mathrm{tr}(D) = \sum_{k=1}^{n} \lambda_k = \lambda_q = \theta > 0$$

    C'est absurde.
  - × Soit $r \in [\![1, n]\!]$ un entier distinct de q et vérifiant $\lambda_r \neq 0$.
    Tous les termes de la somme étant positifs, on a :

$$2m = \sum_{k=1}^{p} \lambda_k^2 \geqslant \lambda_q^2 + \lambda_r^2 > \theta^2$$

$$\boxed{\text{Par stricte croissance de la fonction racine carrée sur } [0, +\infty[\text{, il vient : } \sqrt{2m} > |\theta| = \theta.}$$

□

Commentaire

L'encadrement avec des inégalités larges est beaucoup plus facile à montrer. Cette question est à considérer comme difficile car il faut prendre des initiatives pour obtenir les inégalités strictes demandées.

c) Soit $r \in \mathbb{N}^*$ et $x_1, \ldots, x_r$ des réels. Montrer que : $\sum\limits_{1 \leqslant i,j \leqslant r} 2x_i x_j \leqslant \sum\limits_{1 \leqslant i,j \leqslant r} (x_i^2 + x_j^2)$.

En déduire que $\left(\sum\limits_{i=1}^{r} x_i\right)^2 \leqslant r \left(\sum\limits_{i=1}^{r} x_i^2\right)$.

Démonstration.

- Soit $(i,j) \in [\![1,r]\!]^2$. Comme $(x_i - x_j)^2 \geqslant 0$, $x_i^2 - 2x_i x_j + x_j^2 \geqslant 0$ et donc $2x_i x_j \leqslant x_i^2 + x_j^2$.
  On somme ces inégalités pour tous les couples $(i,j) \in [\![1,r]\!]^2$:

  $$\boxed{\text{On obtient : } \sum_{1 \leqslant i,j \leqslant r} 2x_i x_j \leqslant \sum_{1 \leqslant i,j \leqslant r} (x_i^2 + x_j^2).}$$

- D'autre part :

$$\begin{aligned}
\left(\sum_{i=1}^{r} x_i\right)^2 &= \left(\sum_{i=1}^{r} x_i\right)\left(\sum_{j=1}^{r} x_j\right) \\
&= \sum_{1 \leqslant i,j \leqslant r} x_i x_j \\
&\leqslant \sum_{1 \leqslant i,j \leqslant r} \frac{1}{2}(x_i^2 + x_j^2) && \textit{(d'après ce qui précède)} \\
&= \frac{1}{2}\left(\left(\sum_{1 \leqslant i,j \leqslant r} x_i^2\right) + \left(\sum_{1 \leqslant i,j \leqslant r} x_j^2\right)\right) && \textit{(par linéarité de la somme)} \\
&= \frac{1}{2}\left(\left(\sum_{j=1}^{r}\sum_{i=1}^{r} x_i^2\right) + \left(\sum_{i=1}^{r}\sum_{j=1}^{r} x_j^2\right)\right) \\
&= \frac{1}{2} \times 2 \sum_{j=1}^{r}\sum_{i=1}^{r} x_i^2 && \textit{(par symétrie)} \\
&= r\left(\sum_{i=1}^{r} x_i^2\right) && \textit{(car } \left(\sum_{i=1}^{r} x_i^2\right) \textit{ ne dépend pas de } j\textit{)}
\end{aligned}$$

$$\boxed{\text{On a bien : } \left(\sum_{i=1}^{r} x_i\right)^2 \leqslant r\left(\sum_{i=1}^{r} x_i^2\right).}$$

□

Commentaire

L'égalité $\left(\sum\limits_{i=1}^{r} x_i\right)\left(\sum\limits_{j=1}^{r} x_j\right) = \sum\limits_{1 \leqslant i,j \leqslant r} x_i x_j$ n'a sans doute pas besoin d'être détaillée dans un sujet difficile comme celui-ci, puisque ce n'est pas le cœur de la question et il faut donc la connaître.

$$\begin{aligned}
\left(\sum_{i=1}^{r} x_i\right)\left(\sum_{j=1}^{r} x_j\right) &= \sum_{j=1}^{r}\left(\left(\sum_{i=1}^{r} x_i\right) x_j\right) && \textit{(par linéarité de la somme sur } j\textit{)} \\
&= \sum_{j=1}^{r}\left(\sum_{i=1}^{r} x_i x_j\right) && \textit{(par linéarité de la somme sur } i\textit{)} \\
&= \sum_{1 \leqslant i,j \leqslant r} x_i x_j
\end{aligned}$$

d) En conclure que $\mathcal{E}(A) \leqslant \theta + \sqrt{(p-1)(2m-\theta^2)}$.

Démonstration.
Numérotons $x_1, \dots, x_p$ les réels $|\lambda_1|, \dots, |\lambda_p|$ de telle sorte que $x_p = |\theta|$.
Remarquons alors que, pour tout $k \in [\![1, p]\!]$, $x_k^2 = \lambda_k^2$.
D'après la question ***13.b)***, $|\theta| = \theta$ et

$$2m = \sum_{k=1}^{p} \lambda_k^2 = \sum_{i=1}^{p} x_i^2 = \theta^2 + \sum_{i=1}^{p-1} x_i^2$$

d'où :

$$(p-1)(2m-\theta^2) = (p-1)\sum_{i=1}^{p-1} x_i^2$$

et, d'après la question ***13.c)*** appliquée avec $r = p-1 \geqslant 1$ (car $p \geqslant 2$) :

$$(p-1)\sum_{i=1}^{p-1} x_i^2 \geqslant \left(\sum_{i=1}^{p-1} x_i\right)^2$$

De plus :

$$\begin{aligned} \left(\sum_{i=1}^{p-1} x_i\right)^2 &= \left(\left(\sum_{i=1}^{p} x_i\right) - x_p\right)^2 \\ &= \left(\left(\sum_{k=1}^{p} |\lambda_k|\right) - \theta\right)^2 \\ &= (\mathcal{E}(A) - \theta)^2 \end{aligned}$$

*(car les autres valeurs propres sont nulles d'après la question **13.a)**)*

D'où :

$$(p-1)(2m-\theta^2) \geqslant (\mathcal{E}(A) - \theta)^2$$

On vérifie alors :
× $p-1 \geqslant 0$ car $p \geqslant 2$,
× $2m - \theta^2 = \sum_{k=1}^{p-1} x_k^2 \geqslant 0$,
× $\mathcal{E}(A) - \theta = \sum_{i=1}^{p-1} x_i \geqslant 0$ car tous les x_i sont positifs.

Ainsi, par croissance de la fonction racine carrée sur $[0, +\infty[$:

$$\boxed{\mathcal{E}(A) \leqslant \theta + \sqrt{(p-1)(2m-\theta^2)}.}$$

□

14. On admet qu'on peut choisir la matrice P de la question ***12.a)*** de sorte que $P^{-1} = {}^tP$.
On pose alors $Q = {}^tP$.
Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ une matrice colonne appartenant à $\mathscr{M}_{n,1}(\mathbb{R})$ et $Y = QX = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

- *On admet que si M est une matrice carrée appartenant à $\mathscr{M}_n(\mathbb{R})$ et Z une matrice colonne appartenant à $\mathscr{M}_{n,1}(\mathbb{R})$ alors ${}^t(MZ) = {}^tZ{}^tM$.*
- *Si U et V sont deux matrices colonnes appartenant à $\mathscr{M}_{n,1}(\mathbb{R})$, ${}^tUV$ est une matrice carrée appartenant à $\mathscr{M}_{1,1}(\mathbb{R})$, on l'identifie à son unique coefficient. Donc ${}^tUV \in \mathbb{R}$.*

Commentaire

Une matrice inversible P vérifiant $P^{-1} = {}^tP$ est appelée *matrice orthogonale*.
Cette notion d'orthogonalité provient d'un *produit scalaire* sur l'espace des vecteurs colonnes. Les produits scalaires sont au programme du parcours ECG en maths approfondies, mais pas en maths appliquées.
Au programme de maths appliquées figure le résultat suivant : si $A \in \mathscr{M}_n(\mathbb{R})$ est symétrique (réelle), alors A est diagonalisable. Ce théorème est appelé *théorème spectral* et est en fait plus précis : A est diagonalisable et il existe une base *orthonormée* de $\mathscr{M}_{n,1}(\mathbb{R})$ formée de vecteurs propres de A. On peut le reformuler de la manière suivante : A est semblable à une matrice diagonale via une matrice de passage *orthogonale*.
On démontre dans les questions ***14.a)***, ***14.b)*** et ***14.c)*** plusieurs résultats élémentaires sur les matrices symétriques qui sont une conséquence de ce théorème spectral.

a) Montrer que $Q^{-1} = {}^tQ$ puis que $A = {}^tQDQ$ et ${}^tYY = {}^tXX$.

Démonstration.
Tout d'abord :

$${}^tQQ = {}^t({}^tP)\,{}^tP = P\,{}^tP = PP^{-1} = I_n$$

donc Q est inversible et $Q^{-1} = {}^tQ$.
Ensuite :

$${}^tQDQ = {}^t({}^tP)D\,{}^tP = PDP^{-1} = P(P^{-1}AP)P^{-1} = A \qquad (\text{car } P^{-1}AP = D)$$

Enfin :

$${}^tYY = {}^t(QX)QX = {}^tX{}^tQQX = {}^tXI_nX = {}^tXX$$

$$\boxed{\text{On a bien : } Q^{-1} = {}^tQ,\ A = {}^tQDQ \text{ et } {}^tYY = {}^tXX.}$$

□

b) Montrer que ${}^tXAX = {}^tYDY = \sum\limits_{k=1}^{n} \lambda_k y_k^2$.

Démonstration.
Tout d'abord :

$${}^tYDY = {}^t(QX)DQX = {}^tX{}^tQDQX = {}^tXAX \qquad \textit{(car } {}^tQDQ = A \textit{ d'après la question } \textbf{\textit{14.a)}}\textit{)}$$

Ensuite :

$$DY = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \lambda_1 y_1 \\ \vdots \\ \lambda_n y_n \end{pmatrix}$$

et

$${}^tYDY = \begin{pmatrix} y_1 & \ldots & \ldots & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 y_1 \\ \vdots \\ \lambda_n y_n \end{pmatrix} = \sum_{k=1}^{n} \lambda_k y_k^2 \qquad (\text{en utilisant l'identification de } \mathscr{M}_{1,1}(\mathbb{R}) \text{ à } \mathbb{R})$$

$$\boxed{\text{On a bien : } {}^tXAX = {}^tYDY = \sum_{k=1}^{n} \lambda_k y_k^2.}$$

□

c) En remarquant que $\sum\limits_{k=1}^{n} y_k^2 = {}^tYY$, en déduire que ${}^tXAX \leqslant \theta\ {}^tXX$.

Démonstration.
Vérifions la première égalité :

$$ {}^tYY = \begin{pmatrix} y_1 & \dots\dots & y_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{k=1}^{n} y_k^2 $$

Par définition de θ :

$$ \forall k \in [\![1,n]\!],\ \lambda_k \leqslant \theta $$

et donc,

$$ \forall k \in [\![1,n]\!],\ \lambda_k y_k^2 \leqslant \theta y_k^2 \qquad (\text{car } y_k^2 \geqslant 0) $$

En sommant toutes ces inégalités, on obtient :

$$ \sum_{k=1}^{n} \lambda_k y_k^2 \leqslant \theta \sum_{k=1}^{n} y_k^2 \quad (*) $$

Ainsi,

$$ \begin{aligned} {}^tXAX &= \sum_{k=1}^{n} \lambda_k y_k^2 && \textit{(d'après la question } \textbf{\textit{14.b)}}\textit{)} \\ &\leqslant \theta \sum_{k=1}^{n} y_k^2 && \textit{(d'après } (*)\textit{)} \\ &= \theta\ {}^tYY \\ &= \theta\ {}^tXX && \textit{(d'après la question } \textbf{\textit{14.a)}}\textit{)} \end{aligned} $$

$$ \boxed{\text{On obtient bien : } {}^tXAX \leqslant \theta\ {}^tXX.} $$

□

15. Soit U la matrice colonne appartenant à $\mathscr{M}_{n,1}(\mathbb{R})$ dont le $i$ème coefficient vaut 1 si i n'est pas isolé et 0 sinon.

a) Montrer que ${}^tUAU = \sum\limits_{(i,j)\in I^2} a_{i,j} = 2m$. En déduire que $\dfrac{2m}{p} \leqslant \theta$.

Démonstration.
Notons $U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$. D'après l'énoncé, pour tout $i \in [\![1,n]\!]$, $u_i = \begin{cases} 1 & \text{si } i \in I \\ 0 & \text{sinon} \end{cases}$.
On a alors, pour tout $k \in [\![1,n]\!]$,

$$ \begin{aligned} [AU]_k &= \sum_{j=1}^{n} a_{k,j} u_j \\ &= \sum_{\substack{j=1 \\ j\in I}}^{n} a_{k,j} \underbrace{u_j}_{=1} + \sum_{\substack{j=1 \\ j\notin I}}^{n} a_{k,j} \underbrace{u_j}_{=0} \\ &= \sum_{\substack{j=1 \\ j\in I}}^{n} a_{k,j} \end{aligned} $$

D'où :

$$\begin{aligned}
{}^tUAU &= \sum_{i=1}^{n} [{}^tU]_i [AU]_i \\
&= \sum_{i=1}^{n} \left(u_i \sum_{\substack{j=1 \\ j\in I}}^{n} a_{i,j} \right) \\
&= \sum_{\substack{i=1 \\ i\in I}}^{n} \left(\underbrace{u_i}_{=1} \sum_{\substack{j=1 \\ j\in I}}^{n} a_{i,j} \right) + \sum_{\substack{i=1 \\ i\notin I}}^{n} \left(\underbrace{u_i}_{=0} \sum_{\substack{j=1 \\ j\in I}}^{n} a_{i,j} \right) \\
&= \sum_{\substack{i=1 \\ i\in I}}^{n} \left(\sum_{\substack{j=1 \\ j\in I}}^{n} a_{i,j} \right) \\
&= \sum_{(i,j)\in I^2} a_{i,j} \\
&= 2m && \textit{(d'après la question } \textbf{\textit{11.a)}}\textit{)}
\end{aligned}$$

On applique la question ***14.c)*** avec $X = U$, on obtient alors :

$${}^tUAU \leqslant \theta\ {}^tUU$$

Or,

$${}^tUU = \sum_{k=1}^{n} u_k^2 = \sum_{k\in I} 1 = p$$

On en déduit que $2m \leqslant \theta p$.

$$\boxed{\text{Finalement, puisque } p > 0 \text{, on a bien : } \dfrac{2m}{p} \leqslant \theta.}$$

□

b) Établir que : $\dfrac{1}{\sqrt{p}} \leqslant \dfrac{\sqrt{2m}}{p} \leqslant \dfrac{\theta}{\sqrt{2m}} \leqslant 1$.

Démonstration.

- Démontrons la première inégalité :

$$\begin{aligned}
\frac{1}{\sqrt{p}} \leqslant \frac{\sqrt{2m}}{p} &\iff 1 \leqslant \sqrt{\frac{2m}{p}} \\
&\iff 1 \leqslant \frac{2m}{p}
\end{aligned}$$

  L'inégalité $1 \leqslant \dfrac{2m}{p}$ étant vraie d'après la question ***11.b)***, on en déduit :

$$\boxed{\frac{1}{\sqrt{p}} \leqslant \frac{\sqrt{2m}}{p}}$$

- Démontrons maintenant la deuxième inégalité :

$$\frac{\sqrt{2m}}{p} \leqslant \frac{\theta}{\sqrt{2m}} \iff \frac{2m}{p} \leqslant \theta$$

L'inégalité $\dfrac{2m}{p} \leqslant \theta$ étant vraie d'après la question ***15.a)***, on en déduit :

$$\boxed{\frac{\sqrt{2m}}{p} \leqslant \frac{\theta}{\sqrt{2m}}}$$

- Démontrons enfin la dernière inégalité :

$$\frac{\theta}{\sqrt{2m}} \leqslant 1 \iff \theta \leqslant \sqrt{2m}$$

L'inégalité $\theta \leqslant \sqrt{2m}$ étant vraie d'après la question ***13.b)***, on en déduit :

$$\boxed{\frac{\theta}{\sqrt{2m}} \leqslant 1}$$

□

16. a) Étudier la fonction $F : x \mapsto x + \sqrt{(p-1)(1-x^2)}$ sur $[0,1]$.

Démonstration.
La fonction F est de la forme $F = \mathrm{id}_{\mathbb{R}} + \sqrt{p-1}\sqrt{u}$ où $u : x \mapsto 1 - x^2$ vérifie :
× u est dérivable sur $[0,1]$,
× pour tout $x \in [0,1[$, $u(x) > 0$ et $u(1) = 0$.
De plus, rappelons que la fonction $x \mapsto \sqrt{x}$ est continue sur $[0,+\infty[$ et dérivable sur $]0,+\infty[$.
Ainsi, par somme et composition, la fonction F est continue sur $[0,1]$ et dérivable sur $[0,1[$.
Soit $x \in [0,1[$.

$$\begin{aligned} F'(x) &= 1 + \sqrt{p-1}\frac{-2x}{2\sqrt{1-x^2}} \\ &= 1 - \frac{x\sqrt{p-1}}{\sqrt{1-x^2}} \\ &= \frac{\sqrt{1-x^2} - x\sqrt{p-1}}{\sqrt{1-x^2}} \end{aligned}$$

Comme $\sqrt{1-x^2} > 0$:

$$\begin{aligned} F'(x) > 0 &\iff \sqrt{1-x^2} - x\sqrt{p-1} > 0 \\ &\iff \sqrt{1-x^2} > x\sqrt{p-1} \\ &\iff 1 - x^2 > x^2(p-1) && \textit{(car } x \mapsto x^2 \textit{ est strictement croissante sur } [0,+\infty[\textit{)} \\ &\iff 1 > px^2 \\ &\iff x^2 < \frac{1}{p} && \textit{(car } p > 0\textit{)} \\ &\iff \sqrt{x^2} < \frac{1}{\sqrt{p}} && \textit{(car } x \mapsto \sqrt{x} \textit{ est strictement croissante sur } [0,+\infty[\textit{)} \\ &\iff |x| < \frac{1}{\sqrt{p}} \\ &\iff -\frac{1}{\sqrt{p}} < x < \frac{1}{\sqrt{p}} \\ &\iff 0 \leqslant x < \frac{1}{\sqrt{p}} && \textit{(car } x \geqslant 0 > -\frac{1}{\sqrt{p}}\textit{)} \end{aligned}$$

De plus : $F'(x) = 0 \iff x = \dfrac{1}{\sqrt{p}}$ (car $x \geqslant 0$).

Dressons maintenant le tableau de variations de F sur $[0, 1]$.

x	0		$\frac{1}{\sqrt{p}}$		1
Signe de $F'(x)$		$+$	0	$-$	$\Vert$
Variations de F	$\sqrt{p-1}$	$\nearrow$	$\sqrt{p}$	$\searrow$	1

$\square$

b) En déduire que $\mathcal{E}(A) \leqslant \sqrt{2m}\ F\left(\dfrac{\sqrt{2m}}{p}\right)$, c'est-à-dire que :

$$\mathcal{E}(A) \leqslant \frac{2m}{p} + \frac{1}{p}\sqrt{(p-1)2m(p^2-2m)} \qquad (1)$$

Démonstration.
On a :

$$\begin{aligned}
\mathcal{E}(A) &\leqslant \theta + \sqrt{(p-1)(2m-\theta^2)} && \textit{(d'après la question } \textbf{\textit{13.d)}}\textit{)} \\
&\leqslant \theta + \sqrt{2m(p-1)\left(1-\frac{\theta^2}{2m}\right)} \\
&= \theta + \sqrt{2m}\sqrt{(p-1)\left(1-\left(\frac{\theta}{\sqrt{2m}}\right)^2\right)} \\
&= \sqrt{2m}\left(\frac{\theta}{\sqrt{2m}} + \sqrt{(p-1)\left(1-\left(\frac{\theta}{\sqrt{2m}}\right)^2\right)}\right) \\
&= \sqrt{2m}\ F\left(\frac{\theta}{\sqrt{2m}}\right)
\end{aligned}$$

De plus :

× d'après la question ***16.a)***, la fonction F est décroissante sur $\left[\frac{1}{\sqrt{p}}, 1\right]$,

× d'après la question ***15.b)***, $\dfrac{1}{\sqrt{p}} \leqslant \dfrac{\sqrt{2m}}{p} \leqslant \dfrac{\theta}{\sqrt{2m}} \leqslant 1$.

D'où : $F\left(\dfrac{\theta}{\sqrt{2m}}\right) \leqslant F\left(\dfrac{\sqrt{2m}}{p}\right)$.

En multipliant cette inégalité par $\sqrt{2m} \geqslant 0$ et par transitivité, on obtient :

$$\mathcal{E}(A) \leqslant \sqrt{2m}\ F\left(\frac{\sqrt{2m}}{p}\right)$$

Par ailleurs :

$$\begin{aligned}
\sqrt{2m}\ F\left(\frac{\sqrt{2m}}{p}\right) &= \sqrt{2m}\left(\frac{\sqrt{2m}}{p} + \sqrt{(p-1)\left(1-\left(\frac{\sqrt{2m}}{p}\right)^2\right)}\right) \\
&= \frac{2m}{p} + \sqrt{2m}\ \sqrt{(p-1)\left(1-\frac{2m}{p^2}\right)} \\
&= \frac{2m}{p} + \sqrt{2m}\ \sqrt{(p-1)\ \frac{p^2-2m}{p^2}} \\
&= \frac{2m}{p} + \frac{1}{p}\sqrt{(p-1)2m(p^2-2m)}
\end{aligned}$$

$$\boxed{\text{Finalement, on a bien : } \mathcal{E}(A) \leqslant \frac{2m}{p} + \frac{1}{p}\sqrt{(p-1)2m(p^2-2m)}.}$$

□

17. On suppose dans cette question que A est la matrice d'adjacence d'un graphe complet de sommets $1, \ldots, n$ donc $m = \dfrac{n(n-1)}{2}$ et $p = n$.

a) Représenter la matrice A.

Démonstration.
Dans un graphe complet (supposé ici sans boucle), chaque sommet est relié à tous les autres sommets (et est donc de degré $n-1$).

$$\boxed{\text{Pour un tel graphe, } A = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}.}$$

□

b) Montrer que -1 est une valeur propre de A et que le sous-espace propre associé est de dimension $n-1$.

Démonstration.
Tout d'abord :

$$\mathrm{rg}(A-(-1)I_n) = \mathrm{rg}(A+I_n) = \mathrm{rg}\left(\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 1 \end{pmatrix}\right) = 1$$

En effet, la première colonne est non nulle et toutes les autres colonnes sont égales à la première.
On en déduit que $\mathrm{rg}(A+I_n) < n$ et donc -1 est une valeur propre de A.
Ensuite, d'après le théorème du rang :

$$\dim(\ker(A+I_n)) + \mathrm{rg}(A+I_n) = \dim(E_{-1}(A)) + \mathrm{rg}(A+I_n) = n$$

$$\boxed{\text{On en déduit que } \dim(E_{-1}(A)) = n-1.}$$

□

c) Établir aussi que $n-1$ est une valeur propre de A. En déduire que $\mathcal{E}(A) = 2(n-1)$ et que l'inégalité (1) est alors une égalité.

Démonstration.
On remarque :

$$\begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} n-1 \\ \vdots \\ n-1 \end{pmatrix} = (n-1) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Or, le vecteur $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ est non nul, donc $n-1$ est une valeur propre de A et $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ est un vecteur propre de A associé à la valeur propre $(n-1)$.
On a nécessairement :

$$\dim(E_{-1}(A)) + \dim(E_{n-1}(A)) \leqslant n$$

sinon on pourrait construire (par théorème de concaténation) une famille libre de cardinal strictement plus grand que n dans $\mathscr{M}_{n,1}(\mathbb{R})$, ce qui serait absurde.
Or, $\dim(E_{-1}(A)) = n-1$ d'après la question ***17.b)*** et donc $\dim(E_{n-1}(A)) \leqslant 1$.
Puisque $n-1$ est valeur propre de A, on a aussi $\dim(E_{n-1}(A)) \geqslant 1$.

$$\boxed{\text{Finalement, } \dim(E_{n-1}(A)) = 1 \text{ et } \mathcal{E}(A) = (n-1) \times |-1| + 1 \times |n-1| = 2(n-1).}$$

De plus, comme $m = \dfrac{n(n-1)}{2}$ et $p = n$:

$$\begin{aligned} \frac{2m}{p} + \frac{1}{p}\sqrt{(p-1)2m(p^2-2m)} &= \frac{n(n-1)}{n} + \frac{1}{n}\sqrt{(n-1)n(n-1)(n^2-n(n-1))} \\ &= n-1+\frac{1}{n}\sqrt{(n-1)^2\ n^2(n-(n-1))} \\ &= n-1+\frac{1}{n}\sqrt{(n-1)^2\ n^2} \\ &= n-1+\frac{1}{n}(n-1)n \\ &= 2(n-1) \end{aligned}$$

$$\boxed{\text{Le graphe complet à } n \text{ sommets est un cas d'égalité dans l'inégalité (1).}}$$

□

Commentaire

L'objet de la question ***17*** est de fournir un exemple explicite de graphe où l'inégalité (1) devient une égalité.
Cela démontre que l'inégalité (1) est **optimale**.
En effet, on ne peut pas améliorer le majorant obtenu de l'énergie d'un graphe (c'est-à-dire trouver un nouveau majorant qui soit strictement plus petit que le précédent), sinon cette nouvelle inégalité serait fausse pour le graphe complet à n sommets.
Lorsque l'on démontre une inégalité (ici la majoration (1) de l'énergie d'un graphe), il faut toujours se demander si celle-ci est optimale.

Commentaire

Nous avons utilisé au cours de la preuve l'inégalité :

$$\dim(E_{-1}(A)) + \dim(E_{n-1}(A)) \leqslant n$$

Ce type d'inégalité (sur la somme des dimensions des sous-espaces propres d'une matrice), autrefois présente dans le programme de mathématiques en ECE, a disparu du programme officiel en 2022. Il faut donc la justifier à l'écrit. Nous avons opté dans ce corrigé pour une justification concise faisant apparaître l'ingrédient principal (le théorème de concaténation) et le type de raisonnement employé (par l'absurde).
Pour permettre une bonne compréhension, nous détaillons ci-dessous la preuve de l'inégalité (dans un cadre un peu plus général), tout en spécifiant qu'un tel niveau de détails n'est certainement pas un attendu à l'écrit dans un tel sujet qui demande une forte prise d'initiative.
Supposons que la matrice A admette deux valeurs propres distinctes λ et μ.
Supposons :

$$\dim(E_\lambda(A)) + \dim(E_\mu(A)) > n$$

Posons $p = \dim(E_\lambda(A))$ et $q = \dim(E_\mu(A))$.
Soit $(U_1, \dots, U_p)$ une base de $E_\lambda(A)$ et soit $(V_1, \dots, V_q)$ une base de $E_\mu(A)$.
Les familles $(U_1, \dots, U_p)$ et $(V_1, \dots, V_q)$ sont deux familles libres constituées de vecteurs propres de A, associés à deux valeurs propres distinctes. Par théorème de concaténation, on en déduit que la famille $(U_1, \dots, U_p, V_1, \dots, V_q)$ est également libre. Ainsi, on a construit une famille libre de cardinal $p+q$ dans $\mathscr{M}_{n,1}(\mathbb{R})$. Or : $\dim(\mathscr{M}_{n,1}(\mathbb{R})) = n$ et $p+q > n$. C'est absurde. D'où :

$$\dim(E_\lambda(A)) + \dim(E_\mu(A)) \leqslant n$$

- ▶ *On note* $\alpha = \min\limits_{1 \leqslant k \leqslant p} |\lambda_k|$ *et* $\beta = \max\limits_{1 \leqslant k \leqslant p} |\lambda_k|$. *On note aussi d le degré maximal des sommets du graphe* $G(A)$.

18. Écrire une fonction **Python** `degMax(A)` qui renvoie le maximum des degrés des sommets du graphe $G(A)$, celui-ci étant donné par sa matrice d'adjacence sous la forme du tableau `numpy A`.

Démonstration.

```
1  def degMax(A):
2      n = len(A)
3      d = 0
4      for i in range(n):
5          deg = 0
6          for j in range(n):
7              if A[i,j] == 1: # si j est un voisin de i
8                  deg += 1
9          # deg contient le degré du sommet i
10         if deg > d:
11             d = deg
12     return d
```

Détaillons le fonctionnement de cette fonction **Python**.

- × On parcourt tous les sommets `i` à l'aide de la boucle `for` qui débute en ligne 4.
- × Pour chaque sommet `i`, on calcule son degré à l'aide d'une nouvelle boucle for (lignes 6 à 8). On stocke ce degré dans la variable `deg`.

× La variable `d` est initialisée à 0 et est (potentiellement) mise à jour à chaque nouveau calcul du degré d'un sommet. A chaque tour de boucle, `d` contient le plus grand de tous les degrés qui ont été calculés jusqu'ici. En fin de boucle, puisque on a calculé les degrés de tous les sommets, `d` contient bien le maximum des degrés du graphe $G(A)$.

□

19. a) Montrer que pour tout $j \in \{1, \dots, p\}$, $|\lambda_j|(\alpha+\beta) \geqslant \lambda_j^2 + \alpha\beta$.

Démonstration.
Soit $j \in \{1, \dots, p\}$.

$$\begin{aligned} |\lambda_j|(\alpha+\beta) \geqslant \lambda_j^2 + \alpha\beta &\iff 0 \geqslant |\lambda_j|^2 - |\lambda_j|(\alpha+\beta) + \alpha\beta \\ &\iff 0 \geqslant (|\lambda_j| - \alpha)(|\lambda_j| - \beta) \end{aligned}$$

Or, par définition de α et β : $0 \leqslant \alpha \leqslant |\lambda_j| \leqslant \beta$.
On en déduit : $|\lambda_j| - \alpha \geqslant 0$ et $|\lambda_j| - \beta \leqslant 0$.

$$\boxed{\text{D'où : } |\lambda_j|(\alpha+\beta) \geqslant \lambda_j^2 + \alpha\beta.}$$

□

b) En déduire que $(\alpha+\beta)\mathcal{E}(A) \geqslant \mathrm{tr}(A^2) + p\alpha\beta$, puis que $\mathcal{E}(A) \geqslant \dfrac{2m + p\alpha\beta}{\alpha+\beta}$.

Démonstration.
Sommons les inégalités de la question précédente pour j variant de 1 à p :

$$(\alpha+\beta)\sum_{j=1}^{p} |\lambda_j| \geqslant \sum_{j=1}^{p} \lambda_j^2 + \sum_{j=1}^{p} \alpha\beta = \sum_{j=1}^{p} \lambda_j^2 + p\alpha\beta$$

Or,

$$\begin{aligned} \sum_{j=1}^{p} \lambda_j^2 &= \sum_{j=1}^{n} \lambda_j^2 && \textit{(d'après la question } \textbf{13.a)}\textit{)} \\ &= \mathrm{tr}(A^2) && \textit{(d'après la question } \textbf{5.b)}\textit{)} \end{aligned}$$

et

$$\begin{aligned} \sum_{j=1}^{p} |\lambda_j| &= \sum_{j=1}^{n} |\lambda_j| && \textit{(d'après la question } \textbf{13.a)}\textit{)} \\ &= \mathcal{E}(A) && \textit{(par définition)} \end{aligned}$$

$$\boxed{\text{D'où : } (\alpha+\beta)\mathcal{E}(A) \geqslant \mathrm{tr}(A^2) + p\alpha\beta.}$$

De plus, A^2 et D^2 sont semblables donc, d'après la question ***4.c)*** : $\mathrm{tr}(A^2) = \mathrm{tr}(D^2)$.
Enfin, d'après la question ***12.b)*** : $\mathrm{tr}(D^2) = 2m$.
D'où :

$$(\alpha+\beta)\mathcal{E}(A) \geqslant 2m + p\alpha\beta$$

De plus, $\alpha \geqslant 0$ et $\beta > 0$ car D n'est pas la matrice nulle.

$$\boxed{\text{Donc } \mathcal{E}(A) \geqslant \frac{2m + p\alpha\beta}{\alpha+\beta}.}$$

□

c) Soit a et b deux réels strictement positifs tels que $a \leqslant b^2$. Étudier les variations de $\varphi : t \mapsto \dfrac{a+bt}{b+t}$ sur $\mathbb{R}^+$.

Démonstration.
La fonction φ est dérivable sur $\mathbb{R}^+$ comme quotient de deux fonctions polynomiales dérivables sur $\mathbb{R}^+$ dont le dénominateur ne s'annule pas sur $\mathbb{R}^+$ (car $b > 0$).
Soit $t \in \mathbb{R}^+$.

$$\varphi'(t) = \frac{b(b+t) - (a+bt)}{(b+t)^2} = \frac{b^2 - a}{(b+t)^2} \geqslant 0 \qquad \textit{(car } a \leqslant b^2\textit{)}$$

$$\boxed{\text{Ainsi : la fonction } \varphi \text{ est croissante sur } \mathbb{R}^+.}$$

□

d) Montrer que $2m \leqslant p\beta^2$. En déduire que $\mathcal{E}(A) \geqslant \dfrac{2m}{\beta}$.

Démonstration.

- Tout d'abord :

$$\begin{aligned} 2m &= \mathrm{tr}(D^2) && \textit{(d'après la question } \textbf{\textit{12.b)}}\textit{)} \\ &= \sum_{k=1}^{n} \lambda_k^2 \\ &= \sum_{k=1}^{p} \lambda_k^2 && \textit{(d'après la question } \textbf{\textit{13.a)}}\textit{)} \\ &= \sum_{k=1}^{p} |\lambda_k|^2 \end{aligned}$$

  Ensuite, par définition de β, pour tout $k \in [\![1, p]\!]$, $0 \leqslant |\lambda_k| \leqslant \beta$.
  Par croissance de la fonction $x \mapsto x^2$ sur $\mathbb{R}^+$, pour tout $k \in [\![1, p]\!]$, $|\lambda_k|^2 \leqslant \beta^2$.
  En sommant ces inégalités, on obtient :

$$\sum_{k=1}^{p} |\lambda_k|^2 \leqslant \sum_{k=1}^{p} \beta^2 = p\beta^2$$

$$\boxed{\text{D'où : } 2m \leqslant p\beta^2.}$$

- D'après la question ***19.b)*** :

$$\mathcal{E}(A) \geqslant \frac{2m + p\alpha\beta}{\alpha + \beta}$$

  Or :

$$\begin{aligned} \frac{2m + p\alpha\beta}{\alpha+\beta} &= p\ \frac{\dfrac{2m}{p} + \alpha\beta}{\alpha+\beta} \\ &= p\ \frac{a + b\alpha}{b + \alpha} && \textit{(où } a = \frac{2m}{p} \textit{ et } b = \beta\textit{)} \\ &= p\,\varphi(\alpha) \end{aligned}$$

et d'après la question ***19.c)*** : $\varphi(\alpha) \geqslant \varphi(0) = \dfrac{a}{b} = \dfrac{2m}{p\beta}$.

$$\boxed{\text{D'où : } \mathcal{E}(A) \geqslant \frac{2m}{\beta}.}$$

□

Commentaire

Si l'on souhaite seulement se débarrasser de α (comme c'est le cas dans cette question), on ne peut pas obtenir de meilleure minoration que $\varphi(\alpha) \geqslant \varphi(0)$ car $\alpha = 0$ est tout à fait possible dans ce contexte. En effet, on sait seulement que $\lambda_{p+1} = \cdots = \lambda_n = 0$, mais on ne sait pas si ce sont les seules valeurs propres nulles. Pour le dire autrement, on sait que $\dim(\ker(A)) \geqslant n-p$ mais on ne sait pas si $\dim(\ker(A)) = n-p$. Dans le cas où la dimension est en fait strictement supérieure à $n-p$, on a $\alpha = 0$.

20. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ un vecteur propre pour une valeur propre λ de A telle que $|\lambda| = \beta$.

a) Montrer que pour tout $i \in \{1, \dots, n\}$, $\beta\,|x_i| \leqslant d \max\limits_{1 \leqslant j \leqslant n} |x_j|$.

Démonstration.
Soit $i \in \{1, \dots, n\}$.

Par définition de X, on a $AX = \lambda X$ et donc : $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$.

Ainsi, en identifiant les coefficients en position i :

$$\lambda x_i = \sum_{k=1}^{n} a_{i,k}\, x_k$$

D'où :

$$\begin{aligned}
|\lambda x_i| &= \left| \sum_{k=1}^{n} a_{i,k}\, x_k \right| \\
&\leqslant \sum_{k=1}^{n} |a_{i,k}\, x_k| && \textit{(par inégalité triangulaire)} \\
&= \sum_{k=1}^{n} a_{i,k}\, |x_k| && \textit{(car } a_{i,k} \geqslant 0\textit{)} \\
&\leqslant \sum_{k=1}^{n} a_{i,k} \left(\max_{1 \leqslant j \leqslant n} |x_j| \right) && \textit{(car, pour tout } k \in [\![1,n]\!],\ |x_k| \leqslant \max_{1 \leqslant j \leqslant n} |x_j|\textit{)} \\
&= \left(\max_{1 \leqslant j \leqslant n} |x_j| \right) \sum_{k=1}^{n} a_{i,k} \\
&\leqslant d \left(\max_{1 \leqslant j \leqslant n} |x_j| \right) && \textit{(car } \sum_{k=1}^{n} a_{i,k} = \deg(i) \leqslant d\textit{)}
\end{aligned}$$

De plus, $|\lambda x_i| = |\lambda|\,|x_i| = \beta\,|x_i|$.

$$\boxed{\text{Pour tout } i \in \{1, \dots, n\},\ \beta\,|x_i| \leqslant d \max_{1 \leqslant j \leqslant n} |x_j|.}$$

□

b) En conclure que $\mathcal{E}(A) \geqslant \dfrac{2m}{d} \geqslant \dfrac{2m}{p-1}$ (2).

Démonstration.

- Fixons $i \in \{1, \ldots, n\}$ tel que $|x_i| = \max\limits_{1 \leqslant j \leqslant n} |x_j|$.
  D'après l'inégalité de la question précédente avec cet entier i : $\beta\,|x_i| \leqslant d\,|x_i|$ $(*)$.
  Montrons que $|x_i| > 0$ en raisonnant par l'absurde.
  Supposons : $|x_i| = 0$.
  Alors, pour tout $j \in [\![1, n]\!]$, $0 \leqslant |x_j| \leqslant |x_i| = 0$ et donc $|x_j| = 0$.
  On en déduit que le vecteur propre X est nul. C'est absurde.
  On peut donc diviser $(*)$ par $|x_i|$ et on obtient $\beta \leqslant d$.
  En passant à l'inverse et en mutipliant par $2m$, on obtient :
  $$\boxed{\mathcal{E}(A) \geqslant \frac{2m}{d}}$$
- Montrons maintenant que $d \leqslant p-1$ en raisonnant par l'absurde.
  Supposons : $d \geqslant p$. Alors il existe un sommet qui possède au moins p voisins (nécessairement distincts de lui-même). On compte donc au moins $p+1$ sommets non isolés. Or, par définition de p, il y a exactement p sommets non isolés. C'est absurde.
  D'où, $0 < d \leqslant p-1$.
  En passant à l'inverse, on obtient : $\dfrac{1}{p-1} \leqslant \dfrac{1}{d}$.
  En multipliant par $2m \geqslant 0$, on a enfin :
  $$\boxed{\frac{2m}{d} \geqslant \frac{2m}{p-1}}$$

$$\boxed{\text{Finalement : } \mathcal{E}(A) \geqslant \frac{2m}{d} \geqslant \frac{2m}{p-1}.}$$

□

c) Montrer que l'égalité dans (2) est réalisée pour la matrice A carrée appartenant à $\mathscr{M}_n(\mathbb{R})$ associée au graphe dont l'unique arête est $\{1, 2\}$.

Démonstration.
Écrivons la matrice d'adjacence de ce graphe par blocs :
$$A = \begin{pmatrix} B & 0_{2,n-2} \\ 0_{n-2,2} & 0_{n-2} \end{pmatrix}$$
où
$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
On remarque que les $n-2$ dernières colonnes de A sont nulles et que les deux premières colonnes sont non colinéaires, donc $\operatorname{rg}(A) = 2$ et 0 est valeur propre de A de multiplicité $n-2$ (autrement dit, $\dim(\ker(A)) = n-2$).

Soit $\lambda \in \mathbb{R}$.

$$\mathrm{rg}(A - \lambda I_n) = \mathrm{rg}\left(\begin{pmatrix} -\lambda & 1 & 0 & \cdots & \cdots & 0 \\ 1 & -\lambda & 0 & \cdots & \cdots & 0 \\ 0 & 0 & -\lambda & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\lambda \end{pmatrix}\right)$$

$$= \mathrm{rg}\left(\begin{pmatrix} 1 & -\lambda & 0 & \cdots & \cdots & 0 \\ -\lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & -\lambda & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\lambda \end{pmatrix}\right) \qquad L_1 \leftrightarrow L_2$$

$$= \mathrm{rg}\left(\underbrace{\begin{pmatrix} 1 & -\lambda & 0 & \cdots & \cdots & 0 \\ 0 & 1-\lambda^2 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & -\lambda & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\lambda \end{pmatrix}}_{=M_\lambda}\right) \qquad L_2 \leftarrow L_2 + \lambda L_1$$

On a obtenu la réduite de Gauss (échelonnée), que l'on note M_λ.
Ainsi :

$$\begin{aligned} \lambda \in \mathrm{Sp}(A) &\iff \text{la matrice } A - \lambda I_n \text{ n'est pas inversible} \\ &\iff \mathrm{rg}(A - \lambda I_n) < n \\ &\iff \mathrm{rg}(M_\lambda) < n \\ &\iff \text{la matrice } M_\lambda \text{ n'est pas inversible} \\ &\iff \text{l'un des coefficients diagonaux de } M_\lambda \text{ est nul} \\ &\iff 1 - \lambda^2 = 0 \ \texttt{OU} \ -\lambda = 0 \\ &\iff \lambda \in \{-1, 0, 1\} \end{aligned}$$

et donc :

$$\mathrm{Sp}(A) = \{-1, 0, 1\}$$

Les valeurs propres -1 et 1 sont chacune de multiplicité 1 ($\dim(E_{-1}(A)) = \dim(E_1(A)) = 1$).
D'où : $\mathcal{E}(A) = 1 \times |-1| + 1 \times |1| + (n-2) \times |0| = 2$.
De plus, pour le graphe considéré : $m = 1$ (il y a une unique arête) et $p = 2$ (il y a exactement deux sommets non isolés : ceux qui sont reliés par l'unique arête du graphe).
D'où $\dfrac{2m}{p-1} = \dfrac{2 \times 1}{2-1} = 2$.

Le graphe à n sommets dont l'unique arête est $\{1, 2\}$ est un cas d'égalité dans l'inégalité (2).

□

Commentaire

Nous avons finalement démontré l'encadrement de l'énergie suivant :

$$\frac{2m}{p-1} \leqslant \mathcal{E}(A) \leqslant \frac{2m}{p} + \frac{1}{p}\sqrt{(p-1)2m(p^2-2m)}$$

De plus, nous avons montré que cet encadrement est optimal au sens où il existe un graphe vérifiant le cas d'égalité à gauche et un autre vérifiant le cas d'égalité à droite.
Remarquons que les deux graphes explicités sont radicalement différents.
Le cas d'égalité à droite est obtenu avec le graphe complet d'ordre n, qui est le graphe le plus connecté possible parmi tous les graphes vérifiant les conditions $m \geqslant 1$ et $n \geqslant p \geqslant 2$, c'est-à-dire parmi les graphes possédant au moins une arête (en effet, si il y a une arête elle relie au moins 2 sommets et donc on a nécessairement $p \geqslant 2$).
Le cas d'égalité à gauche est quant à lui obtenu avec le graphe le moins connecté possible parmi tous les graphes possédant au moins une arête.
A l'aune des résultats démontrés dans ce sujet, on peut faire l'hypothèse que l'énergie d'un graphe soit une mesure de son « taux de connection » : plus un graphe est connecté et plus son énergie est grande.
Pour appuyer cette intuition, remarquons pour terminer que l'énergie d'un graphe complètement déconnecté (qui ne possède aucune arête) est nulle. En effet, sa matrice d'adjacence est la matrice nulle et toutes ses valeurs propres sont nulles.

Aide-mémoire Python

On suppose que l'on a exécuté `import numpy as np, numpy.linalg as al` en début de session.

- Si `T` est un tableau `numpy`, `np.shape(T)` renvoie le nombre de lignes et le nombre de colonnes de T, dans cet ordre, sous la forme d'un couple.
- `np.zeros([p,q])` crée un tableau `numpy` à p lignes et q colonnes ne contenant que des 0.
- `np.ones([p,q])` crée un tableau `numpy` à p lignes et q colonnes ne contenant que des 1.
- `np.eye(p)` crée un tableau `numpy` à p lignes et p colonnes ne contenant que des 1 sur sa diagonale et des 0 ailleurs.
- La fonction `al.eigvals`, appliquée à un tableau `numpy` représentant une matrice symétrique A, renvoie le tableau des coefficients d'une matrice diagonale semblable à A.

HEC 2010 - loi exponentielle, loi géométrique, loi de la somme, de la différence, de la valeur absolue, du $\max$, du $\min$, covariance, coefficient de corrélation linéaire, estimation, loi de Gumbel

- Toutes les variables aléatoires qui interviennent dans ce problème sont supposées définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.
- Sous réserve d'existence, on note $\mathbb{E}(X)$ et $\mathbb{V}(X)$ respectivement l'espérance et la variance d'une variable aléatoire X, et $\mathrm{Cov}(X, Y)$ la covariance de deux v.a.r. X et Y.
- Dans les parties I et III, la fonction de répartition et une densité d'une variable aléatoire X à densité sont notées respectivement F_X et f_X.
- **On admet** que les formules donnant l'espérance et la variance d'une somme de variables aléatoires discrètes, ainsi que la définition et les propriétés de la covariance et du coefficient de corrélation linéaire de deux variables aléatoires discrètes, s'appliquent au cas de variables aléatoires à densité.
- Pour n entier supérieur ou égal à 2, on dit que les variables aléatoires à densité X_1, X_2, ..., X_n sont indépendantes si pour tout n-uplet $(x_1, x_2, \dots, x_n)$ de réels, les événements $[X_1 \leqslant x_1]$, $[X_2 \leqslant x_2]$, ..., $[X_n \leqslant x_n]$ sont indépendants.
- L'objet du problème est double. D'une part, montrer certaines analogies entre les lois géométriques et exponentielles, d'autre part mettre en évidence quelques propriétés asymptotiques de variables aléatoires issues de la loi exponentielle.
  La partie II est indépendante de la partie I.
  La partie III est indépendante de la partie II et largement indépendante de la partie I.

Partie I. Loi exponentielle

1. *a)* Rappeler la valeur de $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\, dt$.

Établir pour tout n de $\mathbb{N}^*$ la convergence de l'intégrale $\displaystyle\int_0^{+\infty} t^n \mathrm{e}^{-t}\, dt$.

On pose alors $I_0 = \displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\, dt$ et pour tout n de $\mathbb{N}^*$ $I_n = \displaystyle\int_0^{+\infty} t^n \mathrm{e}^{-t}\, dt$.

b) Soit n un entier de $\mathbb{N}^*$. À l'aide d'une intégration par parties, établir une relation de récurrence entre I_n et I_{n-1}. En déduire la valeur de I_n en fonction de n.

Soit λ un réel strictement positif. Soient X_1 et X_2 deux variables indépendantes de même loi exponentielle de paramètre λ (d'espérance $\frac{1}{\lambda}$).

On pose : $Y = X_1 - X_2$, $T = \max(X_1, X_2)$ et $Z = \min(X_1, X_2)$.

2. Justifier les relations $T + Z = X_1 + X_2$ et $T - Z = |X_1 - X_2| = |Y|$.

3. *a)* Rappeler sans démonstration les valeurs de $\mathbb{V}(X_1)$ et de $\mathbb{P}\big([X_1 \leqslant x]\big)$, pour tout réel x.

b) Calculer $\mathbb{E}(X_1 + X_2)$, $\mathbb{V}(X_1 + X_2)$, $\mathbb{E}(Y)$, $\mathbb{V}(Y)$.

4. Déterminer pour tout réel z, $F_Z(z)$ et $f_Z(z)$. Reconnaître la loi de Z et en déduire $\mathbb{E}(Z)$ et $\mathbb{V}(Z)$.

5. *a)* Montrer que pour tout réel t, on a : $F_T(t) = \begin{cases} \left(1 - \mathrm{e}^{-\lambda t}\right)^2 & \text{si } t \geqslant 0 \\ 0 & \text{si } t < 0 \end{cases}$.

Exprimer pour tout réel t, $f_T(t)$.

b) Justifier l'existence de $\mathbb{E}(T)$ et $\mathbb{V}(T)$. Montrer que $\mathbb{E}(T) = \dfrac{3}{2\lambda}$ et $\mathbb{V}(T) = \dfrac{5}{4\lambda^2}$.
(on pourra utiliser des changements de variables affine)

6. On note r le coefficient de corrélation linéaire de Z et T. Montrer que $r = \dfrac{1}{\sqrt{5}}$.

7. ***a)*** Préciser $Y(\Omega)$ et $|Y|(\Omega)$.

b) Déterminer une densité de la variable aléatoire $-X_2$.

c) Montrer que pour tout réel y, l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f_{X_1}(t)\, f_{-X_2}(y-t)\, dt$ est convergente et qu'elle vaut $\dfrac{\lambda}{2}\,\mathrm{e}^{-\lambda\,|y|}$ (on distinguera les deux cas : $y \geqslant 0$ et $y < 0$).

d) Établir que la fonction $y \mapsto \dfrac{\lambda}{2}\mathrm{e}^{-\lambda\,|y|}$ est une densité de probabilité sur $\mathbb{R}$; on admet que c'est une densité de la variable aléatoire Y.

e) Déterminer pour tout y réel, $f_{|Y|}(y)$. Reconnaître la loi de $|Y| = T - Z$.

Partie II. Loi géométrique

Soit p un réel de $]0,1[$ et $q = 1-p$.
Soit X_1 et X_2 deux variables indépendantes de même loi géométrique de paramètre p (d'espérance $\frac{1}{p}$).
On pose : $Y = X_1 - X_2$, $T = \max(X_1, X_2)$ et $Z = \min(X_1, X_2)$.
On rappelle que $T + Z = X_1 + X_2$ et $T - Z = |X_1 - X_2| = |Y|$.

1. a) Rappeler sans démonstration les valeurs respectives de $\mathbb{V}(X_1)$ et de $\mathbb{P}([X_1 \leqslant k])$, pour tout k de $X_1(\Omega)$.

Démonstration.
Rappelons tout d'abord que $X_1(\Omega) = \mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$, $\mathbb{P}([X_1 = k]) = p\ q^{k-1}$.

$$\boxed{\mathbb{V}(X_1) = \frac{q}{p^2}}$$

$$\boxed{\text{Enfin, pour tout } k \in \mathbb{N}^*,\ \ \mathbb{P}([X_1 \leqslant k]) = 1 - q^k.}$$

Commentaire

- Notons que la valeur de $\mathbb{P}([X_1 \leqslant k])$ (pour $X_1 \hookrightarrow \mathcal{G}(p)$) n'est pas un attendu du programme. Cependant, c'est une propriété très classique de la loi géométrique. Elle doit donc être connue et on se doit de savoir la redémontrer.
- Rappelons maintenant la démonstration. Soit $k \in \mathbb{N}^*$.
  Notons que : $[X_1 \leqslant k] = \bigcup\limits_{i=1}^{k} [X_1 = i]$. On en déduit :

$$\begin{array}{rcll}
\mathbb{P}([X_1 \leqslant k]) & = & \mathbb{P}\left(\bigcup\limits_{i=1}^{k} [X_1 = i]\right) & \\
& = & \sum\limits_{i=1}^{k} \mathbb{P}([X_1 = i]) & \textit{(les événements } [X_1 = i] \\
& & & \textit{étant incompatibles)} \\
& = & \sum\limits_{i=1}^{k} p\ q^{i-1} & \textit{(car } X_1 \hookrightarrow \mathcal{G}(p)\textit{)} \\
& = & p \sum\limits_{i=0}^{k-1} q^i \ = \ p\ \dfrac{1-q^k}{1-q} & \textit{(car } q \neq 1\textit{)} \\
& = & 1 - q^k &
\end{array}$$

□

b) Calculer $\mathbb{E}(X_1 + X_2)$, $\mathbb{V}(X_1 + X_2)$, $\mathbb{E}(X_1 - X_2)$, $\mathbb{V}(X_1 - X_2)$.

Démonstration.
- Les v.a.r. $X_1 + X_2$ et $X_1 - X_2$ admettent une espérance (resp. une variance) car sont des combinaisons linéaires de v.a.r. qui admettent une espérance (resp. une variance).
- Par linéarité de l'espérance :

$$\begin{array}{rcl}
\mathbb{E}(X_1 + X_2) & = & \mathbb{E}(X_1) + \mathbb{E}(X_2) \\
& = & \dfrac{1}{p} + \dfrac{1}{p} \\
& = & \dfrac{2}{p}
\end{array}
\qquad
\begin{array}{rcl}
\mathbb{E}(X_1 - X_2) & = & \mathbb{E}(X_1) - \mathbb{E}(X_2) \\
& = & \dfrac{1}{p} - \dfrac{1}{p} \\
& = & 0
\end{array}$$

$$\boxed{\mathbb{E}(X_1 + X_2) = \frac{2}{p} \ \text{ et } \ \mathbb{E}(X_1 - X_2) = 0}$$

- Les v.a.r. X_1 et X_2 sont indépendantes. Donc, par propriété de la variance :

$$\begin{aligned}\mathbb{V}(X_1 + X_2) &= \mathbb{V}(X_1) + \mathbb{V}(X_2) \\ &= \frac{q}{p^2} + \frac{q}{p^2} \\ &= 2\,\frac{q}{p^2}\end{aligned}$$

D'après le lemme des coalitions, les v.a.r. X_1 et $-X_2$ sont indépendantes.
Donc, par propriété de la variance :

$$\begin{aligned}\mathbb{V}(X_1 - X_2) &= \mathbb{V}(X_1) + \mathbb{V}(-X_2) \\ &= \mathbb{V}(X_1) + (-1)^2\,\mathbb{V}(X_2) \\ &= \mathbb{V}(X_1) + \mathbb{V}(X_2) \ = \ 2\,\frac{q}{p^2}\end{aligned}$$

$$\boxed{\mathbb{V}(X_1 + X_2) = 2\,\frac{q}{p^2} \quad \text{et} \quad \mathbb{V}(X_1 - X_2) = 2\,\frac{q}{p^2}.}$$

Commentaire

Attention à l'erreur classique. Même si X_1 et X_2 sont indépendantes :

$$\mathbb{V}(X_1 - X_2) \ \cancel{=} \ \mathbb{V}(X_1) - \mathbb{V}(X_2)$$

□

c) Établir la relation : $\mathbb{P}([X_1 = X_2]) = \dfrac{p}{1+q}$.

Démonstration.

- Tout d'abord : $[X_1 = X_2] = [X_1 - X_2 = 0]$.
- La famille $([X_2 = k])_{k\in\mathbb{N}^*}$ est un système complet d'événements.
  Ainsi, d'après la formule des probabilités totales :

$$\begin{aligned}\mathbb{P}([X_1 - X_2 = 0]) &= \sum_{k=1}^{+\infty} \mathbb{P}([X_2 = k] \cap [X_1 - X_2 = 0]) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}([X_2 = k] \cap [X_1 - k = 0]) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}([X_2 = k] \cap [X_1 = k]) \\ &= \sum_{k=1}^{+\infty} \mathbb{P}([X_2 = k])\ \mathbb{P}([X_1 = k]) && \textit{(car X_1 et X_2 sont indépendantes)} \\ &= \sum_{k=1}^{+\infty} p\,q^{k-1}\ p\,q^{k-1} \\ &= p^2 \sum_{k=1}^{+\infty} (q^2)^{k-1} \\ &= p^2 \sum_{k=0}^{+\infty} (q^2)^{k} \ = \ p^2\,\frac{1}{1-q^2} && \textit{(car $q^2 \in \,]0,1[$)} \\ &= p^{\cancel{2}}\,\frac{1}{\cancel{(1-q)}(1+q)} \ = \ \frac{p}{1+q}\end{aligned}$$

$$\boxed{\mathbb{P}([X_1 = X_2]) = \frac{p}{1+q}}$$

□

2. a) Montrer que Z suit la loi géométrique de paramètre $1 - q^2$.
En déduire $\mathbb{E}(Z)$, $\mathbb{V}(Z)$ et $\mathbb{E}(T)$.

Démonstration.

- Comme $Z = \min(X_1, X_2)$ et que $X_1(\Omega) = X_2(\Omega) = \mathbb{N}^*$ alors $Z(\Omega) \subset \mathbb{N}^*$.

Soit $k \in \mathbb{N}^*$.

- Remarquons tout d'abord que :

$$[Z > k] \;=\; [\min(X_1, X_2) > k] \;=\; [X_1 > k] \cap [X_2 > k]$$

On en déduit que :

$$\begin{array}{rcll}
\mathbb{P}([Z > k]) & = & \mathbb{P}([X_1 > k] \cap [X_2 > k]) & \\
& = & \mathbb{P}([X_1 > k]) \times \mathbb{P}([X_2 > k]) & \textit{(car } X_1 \textit{ et } X_2 \textit{ sont indépendantes)} \\
& = & (1 - \mathbb{P}([X_1 \leqslant k]) \times (1 - \mathbb{P}([X_2 \leqslant k]) & \\
& = & (\cancel{1} - (\cancel{1} - q^k)) \times (\cancel{1} - (\cancel{1} - q^k)) & \textit{(d'après la question } \textbf{\textit{1.}}\textit{)} \\
& = & q^k \times q^k & \\
& = & \left(q^2\right)^k &
\end{array}$$

Enfin, comme Z est à valeurs dans $\mathbb{N}^*$:

$$\mathbb{P}([Z > 0]) = \mathbb{P}([Z \in \mathbb{N}^*]) = 1$$

$$\boxed{\forall k \in \mathbb{N},\ \mathbb{P}([Z > k]) = \left(q^2\right)^k}$$

- Par ailleurs, comme Z est à valeurs entières :

$$[Z > k - 1] = [Z = k] \cup [Z > k]$$

Les événements $[Z = k]$ et $[Z > k]$ étant incompatibles :

$$\mathbb{P}([Z > k - 1]) = \mathbb{P}([Z = k]) + \mathbb{P}([Z > k])$$

$$\boxed{\text{Ainsi : } \mathbb{P}([Z = k]) = \mathbb{P}([Z > k - 1]) - \mathbb{P}([Z > k]).}$$

- En combinant ces résultats, on obtient :

$$\begin{array}{rcl}
\mathbb{P}([Z = k]) & = & \mathbb{P}([Z > k - 1]) - \mathbb{P}([Z > k]) \\
& = & \left(q^2\right)^{k-1} - \left(q^2\right)^k \\
& = & \left(q^2\right)^{k-1} (1 - q^2) \\
& = & \left(1 - (1 - q^2)\right)^{k-1} (1 - q^2)
\end{array}$$

$$\boxed{\text{On en déduit que } Z \hookrightarrow \mathcal{G}\left(1 - q^2\right).}$$

$$\boxed{\begin{array}{c}\text{Ainsi } Z \text{ admet une espérance et une variance. De plus :} \\ \mathbb{E}(Z) = \dfrac{1}{1 - q^2} \quad \text{et} \quad \mathbb{V}(Z) = \dfrac{1 - (1 - q^2)}{(1 - q^2)^2} = \dfrac{q^2}{(1 - q^2)^2}.\end{array}}$$

- Enfin, comme $T + Z = \max(X_1, X_2) + \min(X_1, X_2) = X_1 + X_2$, alors :

$$T = X_1 + X_2 - Z$$

La v.a.r. T admet une espérance comme somme de v.a.r. qui admettent une espérance. De plus, par linéarité de l'espérance :

$$\begin{array}{rcl} \mathbb{E}(T) & = & \mathbb{E}(X_1 + X_2 - Z) \\ & = & \mathbb{E}(X_1 + X_2) - \mathbb{E}(Z) \\ & = & \dfrac{2}{p} - \dfrac{1}{1-q^2} \\ & = & \dfrac{2}{1-q} - \dfrac{1}{(1-q)(1+q)} \\ & = & \dfrac{2\ (1+q) - 1}{1-q^2} \ = \ \dfrac{1+2q}{1-q^2} \end{array}$$

$$\boxed{\mathbb{E}(T) = \dfrac{1+2q}{1-q^2}}$$

□

b) Soit k un entier de $\mathbb{N}^*$. Justifier l'égalité : $[Z = k] \cup [T = k] = [X_1 = k] \cup [X_2 = k]$.
En déduire la relation suivante : $\mathbb{P}([T = k]) = 2\ \mathbb{P}([X_1 = k]) - \mathbb{P}([Z = k])$.

Démonstration.

- Tout d'abord :

$$\begin{array}{rl} & \omega \text{ réalise } [Z = k] \ \cup \ [T = k] \\ \Leftrightarrow & Z(\omega) = k \ \texttt{OU} \ T(\omega) = k \\ \Leftrightarrow & \max(X_1(\omega), X_2(\omega)) = k \ \texttt{OU} \ \min(X_1(\omega), X_2(\omega)) = k \end{array}$$

Procédons par disjonstion de cas :

× si $\max(X_1(\omega), X_2(\omega)) = X_1(\omega)$:
Dans ce cas, $\min(X_1(\omega), X_2(\omega)) = X_2(\omega)$ et :

$$\begin{array}{rl} & \max(X_1(\omega), X_2(\omega)) = k \ \texttt{OU} \ \min(X_1(\omega), X_2(\omega)) = k \\ \Leftrightarrow & X_2(\omega) = k \ \texttt{OU} \ X_1(\omega) = k \end{array}$$

× si $\max(X_1(\omega), X_2(\omega)) = X_2(\omega)$:
Dans ce cas :

$$\begin{array}{rl} & \max(X_1(\omega), X_2(\omega)) = k \ \texttt{OU} \ \min(X_1(\omega), X_2(\omega)) = k \\ \Leftrightarrow & X_1(\omega) = k \ \texttt{OU} \ X_2(\omega) = k \end{array}$$

On en déduit que :

$$\begin{array}{rl} & \omega \text{ réalise } [Z = k] \ \cup \ [T = k] \\ \Leftrightarrow & X_1(\omega) = k \ \texttt{OU} \ X_2(\omega) = k \\ \Leftrightarrow & \omega \text{ réalise } [X_1 = k] \ \cup \ [X_2 = k] \end{array}$$

$$\boxed{[Z = k] \cup [T = k] = [X_1 = k] \cup [X_2 = k]}$$

- Or :

$$\begin{array}{rcl}\mathbb{P}([Z=k]\cup[T=k]) & = & \mathbb{P}([Z=k])+\mathbb{P}([T=k])-\mathbb{P}([Z=k]\cap[T=k]) \\ & = & \mathbb{P}([Z=k])+\mathbb{P}([T=k])-\mathbb{P}([X_1=k]\cap[X_2=k])\end{array}$$

En effet, on peut démontrer en procédant comme en début de question que :

$$[Z=k]\cap[T=k] \ = \ [X_1=k]\cap[X_2=k]$$

- De même :

$$\begin{array}{rcl}\mathbb{P}([X_1=k]\cup[X_2=k]) & = & \mathbb{P}([X_1=k])+\mathbb{P}([X_2=k])-\mathbb{P}([X_1=k]\cap[X_2=k]) \\ & = & 2\ \mathbb{P}([X_1=k])-\mathbb{P}([X_1=k]\cap[X_2=k])\end{array}$$

En effet : $\mathbb{P}([X_1=k])=\mathbb{P}([X_2=k])$ car X_1 et X_2 suivent la même loi.

- D'après ce qui précède :

$$\begin{array}{rrcl} & \mathbb{P}([Z=k]\ \cup\ [T=k]) & = & \mathbb{P}([X_1=k]\ \cup\ [X_2=k]) \\ \text{d'où} & \mathbb{P}([Z=k])+\mathbb{P}([T=k])-\cancel{\mathbb{P}([X_1=k]\cap[X_2=k])} & = & 2\ \mathbb{P}([X_1=k])-\cancel{\mathbb{P}([X_1=k]\cap[X_2=k])} \\ \text{et} & \mathbb{P}([T=k]) & = & 2\ \mathbb{P}([X_1=k])-\mathbb{P}([Z=k])\end{array}$$

$$\boxed{\mathbb{P}([T=k]) \ = \ 2\ \mathbb{P}([X_1=k])-\mathbb{P}([Z=k])}$$

Commentaire

- Dans l'énoncé, il est demander de « justifier l'égalité » et non de la démontrer. Cette nuance signifie généralement que des points seront attribués même pour une explication avec les mains.
- Il est pratique pour conclure que de dire que les événements $[T=k]$ et $[Z=k]$ (resp. $[X_1=k]$ et $[X_2=k]$) sont incompatibles. Mais on ne peut en aucun cas affirmer une telle chose !
  Il peut exister $\omega\in\Omega$ tel que $\max(X_1(\omega),X_2(\omega))=k$ et $\min(X_1(\omega),X_2(\omega))=k$.
  Cela se produit pour tout $\omega\in\Omega$ tel que : $X_1(\omega)=X_2(\omega)=k$.

□

c) Établir la formule : $\mathbb{V}(T)=\dfrac{q(2q^2+q+2)}{(1-q^2)^2}$.

Démonstration.

- Tout d'abord, comme $X_1(\Omega)=X_2(\Omega)=\mathbb{N}^*$ alors : $T(\Omega)=\big(\max(X_1,X_2)\big)(\Omega)\subset\mathbb{N}^*$.
- La v.a.r. T admet une variance ssi elle admet un moment d'ordre 2. Autrement dit, la v.a.r. T admet une variance ssi la série $\sum\limits_{k\geqslant 1} k^2\ \mathbb{P}([T=k])$ est absolument convergente.
  Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.
- Soit $N\in\mathbb{N}^*$.

$$\begin{array}{rcl}\sum\limits_{k=1}^{N} k^2\ \mathbb{P}([T=k]) & = & \sum\limits_{k=1}^{N} k^2\ \big(2\ \mathbb{P}([X_1=k])-\mathbb{P}([Z=k])\big) \\ & = & 2\ \sum\limits_{k=1}^{N} k^2\ \mathbb{P}([X_1=k])-\sum\limits_{k=1}^{N} k^2\ \mathbb{P}([Z=k])\end{array}$$

Or X_1 et Z admettent un moment d'ordre 2 car admettent une variance.
De plus, comme $X_1(\Omega) = \mathbb{N}^*$ et $Z(\Omega) = \big(\min(X_1, X_2)\big)(\Omega) \subset \mathbb{N}^*$ alors :

$$\mathbb{E}(X_1^2) = \sum_{k=1}^{+\infty} k^2 \ \mathbb{P}([X_1 = k]) \quad \text{et} \quad \mathbb{E}(Z^2) = \sum_{k=1}^{+\infty} k^2 \ \mathbb{P}([Z = k])$$

- On en déduit que la série $\sum\limits_{k\geqslant 1} k^2 \ \mathbb{P}([T = k])$ est convergente et par passage à la limite dans l'égalité précédente :

$$\begin{aligned}
\sum_{k=1}^{+\infty} k^2 \ \mathbb{P}([T = k]) &= 2 \sum_{k=1}^{+\infty} k^2 \ \mathbb{P}([X_1 = k]) - \sum_{k=1}^{+\infty} k^2 \ \mathbb{P}([Z = k]) \\
&= 2 \ \mathbb{E}(X_1^2) - \mathbb{E}(Z^2) \\
&= 2 \ \big(\mathbb{V}(X_1) + (\mathbb{E}(X_1))^2\big) - \big(\mathbb{V}(Z) + (\mathbb{E}(Z))^2\big) && \textit{(par la formule de Kœnig-Huygens)} \\
&= 2 \ \left(\frac{q}{p^2} + \left(\frac{1}{p}\right)^2\right) - \left(\frac{q^2}{(1-q^2)^2} + \left(\frac{1}{1-q^2}\right)^2\right) && \textit{(d'après les questions précédentes)} \\
&= 2 \ \frac{q+1}{p^2} - \frac{q^2+1}{(1-q^2)^2}
\end{aligned}$$

$$\boxed{T \text{ admet un moment d'ordre 2 et } \mathbb{E}(T^2) = 2 \ \frac{q+1}{p^2} - \frac{q^2+1}{(1-q^2)^2}.}$$

- Par la formule de Kœnig-Huygens :

$$\begin{aligned}
\mathbb{V}(T) &= \mathbb{E}(T^2) - \big(\mathbb{E}(T)\big)^2 \\
&= \left(2 \ \frac{q+1}{p^2} - \frac{q^2+1}{(1-q^2)^2}\right) - \left(\frac{1+2q}{1-q^2}\right)^2 \\
&= 2 \ \frac{q+1}{p^2} - \frac{q^2+1}{(1-q^2)^2} - \frac{1+4q+4q^2}{(1-q^2)^2} \\
&= 2 \ \frac{q+1}{p^2} - \frac{2+4q+5q^2}{(1-q^2)^2} \\
&= 2 \ \frac{q+1}{(1-q)^2} - \frac{2+4q+5q^2}{\big((1-q)\ (1+q)\big)^2} \\
&= 2 \ \frac{(q+1)(1+q)^2}{(1-q)^2 \ (1+q)^2} - \frac{2+4q+5q^2}{\big((1-q)\ (1+q)\big)^2} \\
&= \frac{1}{(1-q)^2 \ (1+q)^2} \big(2 \ (1+q)^3 - (2+4q+5q^2)\big) \\
&= \frac{1}{(1-q)^2 \ (1+q)^2} \big(2 \ (1+3q+3q^2+q^3) - (2+4q+5q^2)\big) \\
&= \frac{1}{(1-q)^2 \ (1+q)^2} \big(2q+q^2+2q^3\big) \ = \ \frac{q \ (2+q+2q^2)}{(1-q)^2 \ (1+q)^2}
\end{aligned}$$

$$\boxed{\mathbb{V}(T) = \frac{q \ (2+q+2q^2)}{(1-q)^2 \ (1+q)^2}}$$

Commentaire

- Il faut prendre le réflexe de connaître la formule de Kœnig-Huygens dans les deux sens.
L'écriture :
$$\mathbb{V}(X_1) \ = \ \mathbb{E}(X_1^2) - (\mathbb{E}(X_1))^2$$
fournit l'égalité :
$$\mathbb{E}(X_1^2) \ = \ \mathbb{V}(X_1) + (\mathbb{E}(X_1))^2$$
qui est utilisée dans la démonstration.
- On pouvait aussi calculer $\mathbb{V}(T)$ en s'aidant de l'égalité :
$$T + Z \ = \ X_1 + X_2$$
Comme $T = X_1 + X_2 - Z$ alors T admet un moment d'ordre 2 comme somme de v.a.r. qui admettent un moment d'ordre 2. De la même manière, $T + Z$ admet un moment d'ordre 2.
On a alors :
$$\mathbb{V}(T+Z) \ = \ \mathbb{V}(X_1 + X_2) \ = \ \frac{2\ q}{p^2}$$
D'autre part :
$$\mathbb{V}(T+Z) \ = \ \mathbb{V}(T) + \mathbb{V}(Z) + 2\ \mathrm{Cov}(T, Z)$$
On en déduit donc que :
$$\begin{aligned} \mathbb{V}(T) &= \frac{2\ q}{p^2} - \mathbb{V}(Z) - 2\ \mathrm{Cov}(T,Z) \\ &= \frac{2\ q}{p^2} - \frac{q^2}{(1-q^2)^2} - 2\ \mathrm{Cov}(T,Z) \end{aligned}$$
Par la formule de Kœnig-Huygens :
$$\begin{aligned} \mathrm{Cov}(T,Z) &= \mathbb{E}(TZ) - \mathbb{E}(T)\ \mathbb{E}(Z) \\ &= \mathbb{E}(X_1X_2) - \mathbb{E}(T)\ \mathbb{E}(Z) && \textit{(car } TZ = X_1X_2 \textit{ (*))} \\ &= \mathbb{E}(X_1)\ \mathbb{E}(X_2) - \mathbb{E}(T)\ \mathbb{E}(Z) && \textit{(car } X_1 \textit{ et } X_2 \textit{ sont indépendantes)} \\ &= \frac{1}{p}\ \frac{1}{p} - \frac{1+2q}{1-q^2}\ \frac{1}{1-q^2} \\ &= \frac{1}{p^2} - \frac{1+2q}{(1-q^2)^2} \end{aligned}$$
((*) $TZ = \max(X_1, X_2)\ \min(X_1, X_2) = X_1\ X_2$)
En combinant tous ces résultats, on obtient :
$$\begin{aligned} \mathbb{V}(T) &= \frac{2\ q}{p^2} - \frac{q^2}{(1-q^2)^2} - 2\left(\frac{1}{p^2} - \frac{1+2q}{(1-q^2)^2}\right) \\ &= \frac{2\ (q-1)}{p^2} + \frac{-q^2+4q+2}{(1-q^2)^2} \\ &= \frac{2\ (q-1)}{(1-q)^2} + \frac{-q^2+4q+2}{(1-q^2)^2} \\ &= \frac{1}{(1-q^2)^2}\left(2\ (q-1)(1+q)^2 - q^2 + 4q + 2\right) \\ &= \frac{1}{(1-q^2)^2}\ \left(2q + q^2 + 2q^3\right) \ = \ \frac{q\ (2+q+2q^2)}{(1-q^2)^2} \end{aligned}$$

□

3. ***a)*** Préciser $(T-Z)(\Omega)$. Exprimer pour tout j de $\mathbb{N}^*$, l'évènement $[Z=j]\cap[Z=T]$ en fonction des évènements $[X_1=j]$ et $[X_2=j]$.
En déduire pour tout j de $\mathbb{N}^*$, l'expression de $\mathbb{P}([Z=j]\cap[Z=T])$.

Démonstration.

- Rappelons que : $T-Z=|X_1-X_2|$.
Tout d'abord, comme $X_1(\Omega)=X_2(\Omega)=\mathbb{N}^*$ alors :

$$\begin{array}{rcl}(X_1-X_2)(\Omega) & = & \{X_1(\omega)-X_2(\omega) \mid \omega\in\Omega\} \\ & \subset & \{i-j \mid (i,j)\in(\mathbb{N}^*)^2\} \ = \ \mathbb{Z}\end{array}$$

Ainsi, $(|X_1-X_2|)(\Omega)\subset\mathbb{Z}\cap\mathbb{R}_+=\mathbb{N}$.

$$\boxed{(T-Z)(\Omega)\subset\mathbb{N}}$$

- Soit $j\in\mathbb{N}^*$.

$$\begin{array}{rcl}[Z=j] \ \cap \ [Z=T] & = & [\min(X_1,X_2)=j] \ \cap \ [\min(X_1,X_2)=\max(X_1,X_2)] \\ & = & [\min(X_1,X_2)=j] \ \cap \ [\max(X_1,X_2)=j] \\ & = & [X_1=j] \ \cap \ [X_2=j]\end{array}$$

$$\boxed{\forall j\in\mathbb{N}^*,\ [Z=j] \ \cap \ [Z=T]=[X_1=j] \ \cap \ [X_2=j]}$$

- Ainsi :

$$\begin{array}{rcll}\mathbb{P}([Z=j] \ \cap \ [Z=T]) & = & \mathbb{P}([X_1=j] \ \cap \ [X_2=j]) & \\ & = & \mathbb{P}([X_1=j]) \ \times \ \mathbb{P}([X_2=j]) & \textit{(car } X_1 \textit{ et } X_2 \textit{ sont indépendantes)} \\ & = & p\ q^{j-1}\times p\ q^{j-1} \ = \ p^2\ q^{2j-2} & \end{array}$$

$$\boxed{\forall j\in\mathbb{N}^*,\ \mathbb{P}([Z=j] \ \cap \ [Z=T])=p^2\ q^{2j-2}}$$

□

b) Montrer que pour tout couple (j,l) de $(\mathbb{N}^*)^2$, on a : $\mathbb{P}([Z=j]\cap[T-Z=l])=2p^2q^{2j+l-2}$.

Démonstration.

- Soit $(j,l)\in(\mathbb{N}^*)^2$.

$$\begin{array}{rcl}[Z=j] \ \cap \ [T-Z=l] & = & [Z=j] \ \cap \ [T-j=l] \\ & = & [Z=j] \ \cap \ [T=j+l] \\ & = & [\min(X_1,X_2)=j] \ \cap \ [\max(X_1,X_2)=j+l] \\ & = & \big([X_1=j]\cap[X_2=j+l]\big) \ \cup \ \big([X_1=j+l]\cap[X_2=j]\big)\end{array}$$

Il s'agit d'une réunion de deux événements incompatibles. En effet :

$$\begin{array}{rl} & \big([X_1=j]\cap[X_2=j+l]\big) \ \cap \ \big([X_1=j+l]\cap[X_2=j]\big) \\ = & \big([X_1=j]\cap[X_1=j+l]\big)\cap\big([X_2=j]\cap[X_2=j+l]\big) \\ = & \varnothing \ \cap \ \varnothing \ = \ \varnothing\end{array}$$

En effet : $[X_1=j]\cap[X_1=j+l]=\varnothing$ car $j+l>j$ puisque $l\geqslant 1>0$.

- Ainsi :

$$\begin{aligned}
& \mathbb{P}([Z=j] \ \cap \ [T-Z=l]) \\
= \ & \mathbb{P}\big([X_1=j] \cap [X_2=j+l]\big) \ + \ \mathbb{P}\big([X_1=j+l] \cap [X_2=j]\big) \\
= \ & \mathbb{P}\big([X_1=j]\big) \times \mathbb{P}([X_2=j+l]\big) \ + \ \mathbb{P}\big([X_1=j+l]\big) \times \mathbb{P}\big([X_2=j]\big) && \textit{(car } X_1 \textit{ et } X_2 \textit{ sont indépdendantes)} \\
= \ & p\ q^{j-1} \times p\ q^{j+l-1} \ + \ p\ q^{j+l-1} \times p\ q^{j-1} \\
= \ & p^2\ q^{2j+l-2} + p^2\ q^{2j+l-2} \ = \ 2\ p^2\ q^{2j+l-2}
\end{aligned}$$

$$\boxed{\forall (j,l) \in (\mathbb{N}^*)^2,\ \mathbb{P}([Z=j] \ \cap \ [T-Z=l]) = 2\ p^2\ q^{2j+l-2}}$$

□

c) Montrer que pour tout k de $\mathbb{Z}$, $\mathbb{P}([X_1-X_2=k]) = \dfrac{pq^{|k|}}{1+q}$.
(on distinguera trois cas : $k=0$, $k>0$ et $k<0$)

Démonstration.
Soit $k \in \mathbb{Z}$. On procède par disjonction de cas.

- Si $k=0$:

$$\mathbb{P}([X_1-X_2=k]) = \mathbb{P}([X_1-X_2=0]) = \mathbb{P}([X_1=X_2]) = \frac{p}{1+q} = \frac{pq^0}{1+q}$$

- Si $k>0$:
La famille $([X_2=i])_{i\in\mathbb{N}^*}$ est un système complet d'événements.
Ainsi, par la formule des probabilités totales :

$$\begin{aligned}
\mathbb{P}([X_1-X_2=k]) \ &= \ \sum_{i=1}^{+\infty} \mathbb{P}([X_2=i] \ \cap \ [X_1-X_2=k]) \\
&= \ \sum_{i=1}^{+\infty} \mathbb{P}([X_2=i] \ \cap \ [X_1-i=k]) \\
&= \ \sum_{i=1}^{+\infty} \mathbb{P}([X_2=i] \ \cap \ [X_1=i+k]) \\
&= \ \sum_{i=1}^{+\infty} \mathbb{P}([X_2=i]) \ \times \ \mathbb{P}([X_1=i+k]) && \textit{(car } X_1 \textit{ et } X_2 \textit{ sont indépendantes)} \\
&= \ \sum_{i=1}^{+\infty} p\ q^{i-1} \ \times \ p\ q^{i+k-1} \\
&= \ p^2\ q^k \sum_{i=1}^{+\infty} q^{2i-2} \ = \ p^2\ q^k \sum_{i=1}^{+\infty} \left(q^2\right)^{i-1} \\
&= \ p^2\ q^k \sum_{i=0}^{+\infty} \left(q^2\right)^{i} && \textit{(par décalage d'indice)} \\
&= \ p^2\ q^k\ \frac{1}{1-q^2} \ = \ p^{\cancel{2}}\ q^k\ \frac{1}{\cancel{(1-q)}\ (1+q)} && \textit{(car } q^2 \in \ ]-1,1[\textit{)} \\
&= \ \frac{p\ q^k}{1+q} \ = \ \frac{p\ q^{|k|}}{1+q}
\end{aligned}$$

- Si $k < 0$:

  On procède de la même manière que dans le point précédent. On obtient :

$$\begin{aligned}
\mathbb{P}([X_1 - X_2 = k]) &= \sum_{i=1}^{+\infty} \mathbb{P}([X_2 = i]) \times \mathbb{P}([X_1 = i+k]) \\
&= \sum_{\substack{i=1 \\ \text{tel que } i+k \in \mathbb{N}^*}}^{+\infty} \mathbb{P}([X_2 = i]) \times \mathbb{P}([X_1 = i+k]) \\
&\quad + \sum_{\substack{i=1 \\ \text{tel que } i+k \notin \mathbb{N}^*}}^{+\infty} \mathbb{P}([X_2 = i]) \times \cancel{\mathbb{P}([X_1 = i+k])} && \textit{([}X_1 = i+k\textit{]} = \varnothing \textit{ car } i+k \notin \mathbb{N}^*\textit{)} \\
&= \sum_{i=-k+1}^{+\infty} \mathbb{P}([X_2 = i]) \times \mathbb{P}([X_1 = i+k]) && \left(\textit{car} \begin{array}{cl} & i+k \geqslant 1 \\ \Leftrightarrow & i \geqslant -k+1 \end{array} \right) \\
&= \sum_{i=-k+1}^{+\infty} p\ q^{i-1} \times p\ q^{i+k-1} \\
&= \sum_{i=1}^{+\infty} p\ q^{i-k-1} \times p\ q^{i-1} && \textit{(par décalage d'indice)} \\
&= p^2\ q^{-k} \sum_{i=1}^{+\infty} q^{2i-2} \\
&= \frac{p^2\ q^{-k}}{1+q} = \frac{p\ q^{|k|}}{1+q} && \textit{(en procédant comme au point précédent)}
\end{aligned}$$

$$\boxed{\forall k \in \mathbb{Z},\ \mathbb{P}([X_1 - X_2 = k]) = \frac{p\ q^{|k|}}{1+q}}$$

□

d) En déduire la loi de la variable aléatoire $|X_1 - X_2|$.

Démonstration.

- Rappelons que : $T - Z = |X_1 - X_2|$.

$$\boxed{(|X_1 - X_2|)(\Omega) \subset \mathbb{N}}$$

- Soit $k \in \mathbb{N}$. Remarquons tout d'abord que :

$$[\ |X_1 - X_2| = k] = [X_1 - X_2 = k] \cup [X_1 - X_2 = -k]$$

  Deux cas se présentent :

  × si $k \neq 0$: alors les événements $[X_1 - X_2 = k]$ et $[X_1 - X_2 = -k]$ sont incompatibles. Ainsi :

$$\begin{aligned}
\mathbb{P}([\ |X_1 - X_2| = k]) &= \mathbb{P}([X_1 - X_2 = k]) + \mathbb{P}([X_1 - X_2 = -k]) \\
&= \frac{p\ q^k}{1+q} + \frac{p\ q^k}{1+q} = 2\ \frac{p\ q^k}{1+q} && \textit{(car } |k| = k\textit{)}
\end{aligned}$$

  × si $k = 0$:

$$\mathbb{P}([\ |X_1 - X_2| = 0]) = \mathbb{P}([X_1 - X_2 = 0]) = \frac{p}{1+q}$$

$$\boxed{\forall k \in \mathbb{N}^*, \mathbb{P}([\ |X_1 - X_2| = k]) = 2\ \frac{p\ q^k}{1+q} \quad \text{et} \quad \mathbb{P}([\ |X_1 - X_2| = 0]) = \frac{p}{1+q}}$$

□

e) Établir à l'aide des questions précédentes que les variables Z et $T - Z$ sont indépendantes.

Démonstration.
Il s'agit de démontrer :

$$\forall j \in Z(\Omega),\ \forall l \in (T-Z)(\Omega),\ \mathbb{P}([Z=j] \cap [T-Z=l]) \ = \ \mathbb{P}([Z=j]) \times \mathbb{P}([T-Z=l])$$

avec $Z(\Omega) \subset \mathbb{N}^*$ et $(T-Z)(\Omega) \subset \mathbb{N}$.

- Soit $(j,l) \in (\mathbb{N}^*)^2$.

$$\begin{array}{rll} & \mathbb{P}([Z=j]) \times \mathbb{P}([T-Z=l]) & \\ = & \left(q^2\right)^{j-1} (1-q^2) \times 2 \ \dfrac{p\ q^l}{1+q} & \textit{(d'après les questions précédentes)} \\ = & q^{2j-2}\ (1-q)\ \cancel{(1+q)} \times 2 \ \dfrac{p\ q^l}{\cancel{1+q}} & \\ = & q^{2j-2}\ p \times 2p\ q^l \ = \ 2\ p^2\ q^{2j+l-2} & \\ = & \mathbb{P}([Z=j] \ \cap \ [T-Z=l]) & \textit{(d'après la }\textbf{\textit{3.c)}}\textit{)} \end{array}$$

- Il reste à étudier le cas où $j \in \mathbb{N}^*$ et $l = 0$.

$$\begin{array}{rll} & \mathbb{P}([Z=j]) \times \mathbb{P}([T-Z=0]) & \\ = & \left(q^2\right)^{j-1} (1-q^2) \times \dfrac{p}{1+q} & \textit{(d'après les questions précédentes)} \\ = & q^{2j-2}\ (1-q)\ \cancel{(1+q)} \times \ \dfrac{p}{\cancel{1+q}} & \\ = & q^{2j-2}\ p \times p \ = \ p^2\ q^{2j-2} & \\ = & \mathbb{P}([Z=j] \ \cap \ [T-Z=l]) & \textit{(d'après la }\textbf{\textit{3.a)}}\textit{)} \end{array}$$

$$\boxed{\text{Les variables } Z \text{ et } T-Z \text{ sont indépendantes.}}$$

□

4. a) À l'aide du résultat de la question ***3.e)***, calculer $\text{Cov}(Z,T)$.
Les variables Z et T sont-elles indépendantes ?

Démonstration.

- D'après la question ***3.e)***, les variables Z et $T-Z$ sont indépendantes. On en déduit que :

$$\text{Cov}(Z, T-Z) = 0$$

Or, par linéarité gauche de l'opérateur Cov :

$$\underbrace{\text{Cov}(Z, T-Z)}_{0} \ = \ \text{Cov}(Z,T) \ - \ \underbrace{\text{Cov}(Z,Z)}_{\mathbb{V}(Z)}$$

$$\boxed{\text{Ainsi : } \text{Cov}(Z,T) \ = \ \mathbb{V}(Z) \ = \ \dfrac{q^2}{(1-q^2)^2}}$$

- Comme $q \neq 0$, $\text{Cov}(Z,T) \neq 0$.

$$\boxed{\text{Ainsi, les v.a.r. } Z \text{ et } T \text{ ne sont pas indépendantes.}}$$

□

b) Calculer en fonction de q, le coefficient de corrélation linéaire ρ de Z et T.

Démonstration.
Par définition :

$$\begin{aligned}
\rho(Z,T) &= \frac{\mathrm{Cov}(Z,T)}{\sqrt{\mathbb{V}(Z)}\ \sqrt{\mathbb{V}(T)}} \\
&= \frac{\mathbb{V}(Z)}{\sqrt{\mathbb{V}(Z)}\ \sqrt{\mathbb{V}(T)}} \ = \ \frac{\cancel{\sqrt{\mathbb{V}(Z)}}\ \sqrt{\mathbb{V}(Z)}}{\cancel{\sqrt{\mathbb{V}(Z)}}\ \sqrt{\mathbb{V}(T)}} \\
&= \sqrt{\frac{\frac{q^2}{(1-q^2)^2}}{\frac{q\ (2+q+2q^2)}{(1-q^2)^2}}} \ = \ \sqrt{\frac{q^2}{\cancel{(1-q^2)^2}}\ \frac{\cancel{(1-q^2)^2}}{q\ (2+q+2q^2)}} \\
&= \sqrt{\frac{q}{2+q+2q^2}}
\end{aligned}$$

$$\boxed{\rho(Z,T) = \sqrt{\frac{q}{2+q+2q^2}}}$$

□

c) Déterminer la loi de probabilité du couple (Z,T).

Démonstration.

- On rappelle que $Z(\Omega) \subset \mathbb{N}^*$ et $T(\Omega) \subset \mathbb{N}^*$.
- Soit $(i,j) \in (\mathbb{N}^*)^2$. On procède par disjonction de cas :
  × si $i < j$: alors $[Z = j] \ \cap \ [T = i] = \varnothing$.
  En effet, $Z = \min(X_1, X_2) \ \leqslant \ \max(X_1, X_2) = T$.

$$\boxed{\text{Si } i < j, \ \mathbb{P}([Z = j] \ \cap \ [T = i]) = \mathbb{P}(\varnothing) = 0}$$

  × si $i = j$: alors $[Z = j] \ \cap \ [T = j] = [X_1 = j] \ \cap \ [X_2 = j]$.

$$\begin{aligned}
\mathbb{P}([Z = j] \ \cap \ [T = j]) &= \mathbb{P}([X_1 = j] \ \cap \ [X_2 = j]) \\
&= \mathbb{P}([X_1 = j]) \times \mathbb{P}([X_2 = j]) && \textit{(car } X_1 \textit{ et } X_2 \textit{ sont indépendantes)} \\
&= p\ q^{j-1} \times p\ q^{j-1} \ = \ p^2\ q^{2j-2}
\end{aligned}$$

$$\boxed{\mathbb{P}([Z = j] \ \cap \ [T = j]) \ = \ p^2\ q^{2j-2}}$$

  × si $i > j$: alors $[Z = j] \ \cap \ [T = i] = ([X_1 = j] \cap [X_2 = i]) \ \cup \ ([X_1 = i] \cap [X_2 = j])$.

$$\begin{aligned}
& \mathbb{P}([Z = j] \ \cap \ [T = i]) \\
= \ & \mathbb{P}\big(\ ([X_1 = j] \cap [X_2 = i]) \ \cup \ ([X_1 = i] \cap [X_2 = j]) \ \big) \\
= \ & \mathbb{P}([X_1 = j] \cap [X_2 = i]) \ + \ \mathbb{P}([X_1 = i] \cap [X_2 = j]) && \textit{(car } [X_1 = j] \cap [X_2 = i] \textit{ et } [X_1 = i] \cap [X_2 = j] \textit{ sont incompatibles)} \\
= \ & \mathbb{P}([X_1 = j]) \times \mathbb{P}([X_2 = i]) + \mathbb{P}([X_1 = i]) \times \mathbb{P}([X_2 = j]) && \textit{(car } X_1 \textit{ et } X_2 \textit{ sont indépendantes)} \\
= \ & p\ q^{j-1} \times p\ q^{i-1} + p\ q^{i-1} \times p\ q^{j-1} \ = \ 2\ p^2\ q^{i+j-2}
\end{aligned}$$

$$\boxed{\text{Si } i > j, \ \mathbb{P}([Z = j] \ \cap \ [T = i]) \ = \ 2\ p^2\ q^{i+j-2}.}$$

□

d) Déterminer pour tout j de $\mathbb{N}^*$, la loi de probabilité conditionnelle de T sachant l'évènement $[Z = j]$.

Démonstration.
Soit $j \in \mathbb{N}^*$.

- Rappelons que $T(\Omega) \subset \mathbb{N}^*$.
- Soit $i \in \mathbb{N}^*$.

$$\mathbb{P}_{[Z=j]}([T = i]) = \frac{\mathbb{P}([Z = j] \cap [T = i])}{\mathbb{P}([Z = j])} = \frac{\mathbb{P}([Z = j] \cap [T = i])}{\left(q^2\right)^{j-1} (1 - q^2)}$$

Le loi du couple (Z, T) était donnée par cas, on procède par disjonction de cas :

× si $i < j$: alors $\mathbb{P}([Z = j] \cap [T = i]) = \mathbb{P}(\varnothing) = 0$.

$$\boxed{\text{Si } i < j, \ \mathbb{P}_{[Z=j]}([T = i]) = 0.}$$

× si $i = j$: alors $\mathbb{P}([Z = j] \cap [T = j]) = p^2 \ q^{2j-2}$.

$$\mathbb{P}_{[Z=j]}([T = j]) = \frac{\mathbb{P}([Z = j] \cap [T = j])}{\left(q^2\right)^{j-1} (1 - q^2)} = \frac{p^2 \ \cancel{q^{2j-2}}}{\cancel{\left(q^2\right)^{j-1}} (1 - q^2)} = \frac{p^{\cancel{2}}}{\cancel{(1 - q)} \ (1 + q)} = \frac{p}{1 + q}$$

$$\boxed{\mathbb{P}_{[Z=j]}([T = j]) = \frac{p}{1 + q}}$$

× si $i > j$: alors $\mathbb{P}([Z = j] \cap [T = i]) = 2 \ p^2 q^{i+j-2}$.

$$\begin{aligned}
\mathbb{P}_{[Z=j]}([T = i]) &= \frac{\mathbb{P}([Z = j] \cap [T = j])}{\left(q^2\right)^{j-1} (1 - q^2)} \\
&= \frac{2 \ p^2 \ q^{i+j-2}}{\left(q^2\right)^{j-1} (1 - q^2)} \\
&= \frac{2 \ p^2 \cancel{q^{j-2}} \ q^i}{q^j \ \cancel{q^{j-2}} \ (1 - q^2)} \\
&= \frac{2 \ p^{\cancel{2}} \ q^i}{q^j \ \cancel{(1 - q)} \ (1 + q)} \\
&= \frac{2 \ p \ q^i}{q^j (1 + q)} = \frac{2 \ p \ q^{i-j}}{1 + q}
\end{aligned}$$

$$\boxed{\text{Si } i > j, \ \mathbb{P}_{[Z=j]}([T = i]) = \frac{2 \ p \ q^{i-j}}{1 + q}.}$$

□

e) Soit j un élément de $\mathbb{N}^*$. On suppose qu'il existe une variable aléatoire D_j à valeur dans $\mathbb{N}^*$, dont la loi de probabilité est la loi conditionnelle de T sachant l'évènement $[Z = j]$.
Calculer $\mathbb{E}(D_j)$.

Démonstration.

- La v.a.r. D_j admet une espérance ssi la série $\sum\limits_{i \geqslant 1} i \ \mathbb{P}_{[Z=j]}([T = i])$ est absolument convergente.
Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.
- Soit $N \in \mathbb{N}^*$.

$$\begin{array}{rl} & \sum\limits_{i=1}^{N} i \ \mathbb{P}_{[Z=j]}([T = i]) \\ = & \sum\limits_{i=1}^{j-1} i \ \mathbb{P}_{[Z=j]}([T = i]) \ + \ \sum\limits_{i=j}^{j} i \ \mathbb{P}_{[Z=j]}([T = i]) \ + \ \sum\limits_{i=j+1}^{N} i \ \mathbb{P}_{[Z=j]}([T = i]) \quad \textit{(par relation de Chasles en supposant } N > j\textit{)} \\ = & 0 \ + \ \dfrac{p}{1+q} \ + \ \sum\limits_{i=j+1}^{N} i \ \dfrac{2 \ p \ q^{i-j}}{1+q} \quad \textit{(d'après la question précédente)} \end{array}$$

- Étudions en particulier la somme de droite :

$$\begin{aligned} \sum_{i=j+1}^{N} i \ \frac{2 \ p \ q^{i-j}}{1+q} &= \frac{2 \ p}{1+q} \sum_{i=j+1}^{N} i \ q^{i-j} \\ &= \frac{2 \ p}{1+q} \sum_{i=1}^{N-j} i \ q^{(i+j)-j} \quad \textit{(par décalage d'indice)} \\ &= \frac{2 \ p}{1+q} \sum_{i=1}^{N-j} i \ q^{i} \\ &= \frac{2 \ p \ q}{1+q} \sum_{i=1}^{N-j} i \ q^{i-1} \end{aligned}$$

On reconnaît la somme partielle (d'ordre $N - j$) d'une série géométrique dérivée première convergente car de raison $q \in \,] - 1, 1[$.

- Ainsi, la série $\sum\limits_{i \geqslant 1} i \ \mathbb{P}_{[Z=j]}([T = i])$ est convergente et par passage à la limite dans l'égalité précédente on obtient :

$$\begin{aligned} \sum_{i=1}^{+\infty} i \ \mathbb{P}_{[Z=j]}([T = i]) &= \frac{p}{1+q} + \frac{2 \ p \ q}{1+q} \sum_{i=1}^{+\infty} i \ q^{i-1} \\ &= \frac{p}{1+q} + \frac{2 \ \cancel{p} \ q}{1+q} \ \frac{1}{(1-q)^{\cancel{2}}} \\ &= \frac{1}{(1+q)(1-q)} \left(p \ (1-q) + 2 \ q\right) \\ &= \frac{1}{1-q^2} \left((1-q)^2 + 2 \ q\right) \ = \ \frac{1+q^2}{1-q^2} \end{aligned}$$

$$\boxed{\text{La variable } D_j \text{ admet une espérance et } \mathbb{E}(D_j) = \dfrac{1+q^2}{1-q^2}}$$

Commentaire

- Dans cette question, on détermine $\sum\limits_{i=1}^{+\infty} i\ \mathbb{P}_{[Z=j]}([T=i])$.
  Cette écriture est très proche de l'écriture de $\mathbb{E}(T)$: on a simplement remplacé ici l'application probabilité $\mathbb{P}$ par l'application probabilité $\mathbb{P}_{[Z=j]}$.
  Autrement dit, on détermine l'espérance de T sachant que l'événement $[Z=j]$ est réalisé.
- Cet objet est classique en mathématiques (mais hors programme !). Il s'agit de l'espérance conditionnelle de la variable T relativement à l'événement $[Z=j]$. Elle se note :
  $$\mathbb{E}(T \mid [Z=j])$$
  (cette notation n'est pas très heureuse au vu de la notation utilisée pour noter les probabilités conditionnelles)
- On peut noter que ces calculs d'espérances conditionnelles permettent de déterminer l'espérance. Sous réserve d'existence des objets considérés :
  $$\mathbb{E}(T) \ = \ \sum_{j=1}^{+\infty} \mathbb{P}([Z=j])\ \mathbb{E}(T \mid [Z=j])$$
  Il faut considérer cette égalité comme une formule des probabilités totales (qui est à la base de ce résultat) adaptée à la notion d'espérance.

□

Partie III. Convergences

Dans les questions 1 à 4, λ désigne un paramètre réel strictement positif, inconnu.
Pour n élément de $\mathbb{N}^*$, on considère un n-échantillon $(X_1, X_2, \dots, X_n)$ de variables aléatoires à valeurs strictement positives, indépendantes, de même loi exponentielle de paramètre λ.
On pose pour tout n de $\mathbb{N}^*$: $S_n = \sum_{k=1}^{n} X_k$ et $J_n = \lambda S_n$.

1. Calculer pour tout n de $\mathbb{N}^*$, $\mathbb{E}(S_n)$, $\mathbb{V}(S_n)$, $\mathbb{E}(J_n)$ et $\mathbb{V}(J_n)$.

2. On admet qu'une densité f_{J_n} de J_n est donnée par $f_{J_n}(x) = \begin{cases} \dfrac{\mathrm{e}^{-x}x^{n-1}}{(n-1)!} & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$.

a) À l'aide du théorème de transfert, établir pour tout n supérieur ou égal à 3, l'existence de $\mathbb{E}\left(\dfrac{1}{J_n}\right)$ et de $\mathbb{E}\left(\dfrac{1}{J_n^2}\right)$, et donner leur valeurs respectives.

b) On pose pour tout n supérieur ou égal à 3 : $\widehat{\lambda_n} = \dfrac{n}{S_n}$. Justifier que $\widehat{\lambda_n}$ est un estimateur de λ. Est-il sans biais ? Calculer la limite, lorsque n tend vers $+\infty$, du risque quadratique associé à $\widehat{\lambda_n}$ en λ.

3. Dans cette question, on veut déterminer un intervalle de confiance du paramètre λ au risque α. On note Φ la fonction de répartition de la loi normale centrée réduite, et u_α le réel strictement positif tel que $\Phi(u_\alpha) = 1 - \dfrac{\alpha}{2}$.

a) Énoncer le théorème de la limite centrée. En déduire que la variable aléatoire N_n définie par $N_n = \lambda \dfrac{S_n}{\sqrt{n}} - \sqrt{n}$ converge en loi vers la loi normale centrée réduite.

b) En déduire que pour n assez grand, on a approximativement : $\mathbb{P}\big([-u_\alpha \leqslant N_n \leqslant u_\alpha]\big) = 1 - \alpha$.

c) Montrer que pour n assez grand, l'intervalle $\left[\left(1 - \dfrac{u_\alpha}{\sqrt{n}}\right)\widehat{\lambda_n}, \left(1 + \dfrac{u_\alpha}{\sqrt{n}}\right)\widehat{\lambda_n}\right]$ est un intervalle de confiance de λ au risque α. On note λ_0 la réalisation de $\widehat{\lambda_n}$ sur le n-échantillon.

4. Avec le n-échantillon $(X_1, X_2, \dots, X_n)$, on construit un nouvel intervalle de confiance de λ au risque β ($\beta \neq \alpha$), tel que la longueur de cet intervalle soit k ($k > 1$) fois plus petite que celle obtenue avec le risque α.

a) Justifier l'existence de la fonction réciproque Φ^{-1} de Φ.
Quel est le domaine de définition de Φ^{-1} ?

b) Établir l'égalité $\beta = 2\Phi\left(\dfrac{1}{k}\Phi^{-1}\left(\dfrac{\alpha}{2}\right)\right)$.
En déduire que $\beta > \alpha$. Ce dernier résultat était-il prévisible ?

Dans les questions 5 à 7, on suppose que $\lambda = 1$.

5. On pose pour tout n de $\mathbb{N}^*$: $T_n = \max(X_1, X_2, \dots, X_n)$.
Pour tout n de $\mathbb{N}^*$, pour tout réel x positif ou nul, on pose :

$$g_n(x) = \int_0^x F_{T_n}(t)\ dt \quad \text{et} \quad h_n(x) = \int_0^x t\, f_{T_n}(t)\ dt$$

a) Exprimer $h_n(x)$ en fonction de $F_{T_n}(x)$ et $g_n(x)$.

b) Déterminer pour tout réel t, l'expression de $F_{T_n}(t)$ en fonction de t.
Établir pour tout n supérieur ou égal à 2, la relation : $g_{n-1}(x) - g_n(x) = \dfrac{1}{n}\ F_{T_n}(x)$.

c) En déduire que pour tout n de $\mathbb{N}^*$, pour tout réel x positif ou nul, l'expression de $g_n(x)$ en fonction de x, $F_{T_1}(x)$, $F_{T_2}(x)$, $\dots$, $F_{T_n}(x)$.

d) Montrer que $F_{T_n}(x) - 1$ est équivalent à $-n\mathrm{e}^{-x}$, lorsque x tend vers $+\infty$.

e) Déduire des questions ***c)*** et ***d)*** l'existence de $\mathbb{E}(T_n)$ et montrer que $\mathbb{E}(T_n) = \sum\limits_{k=1}^{n} \dfrac{1}{k}$.

6. On veut étudier dans cette question la convergence en loi de la suite de variables aléatoires $(G_n)_{n\geqslant 1}$ définie par : pour tout n de $\mathbb{N}^*$, $G_n = T_n - \mathbb{E}(T_n)$.
On pose pour tout n de $\mathbb{N}^*$: $\gamma_n = -\ln(n) + \mathbb{E}(T_n)$ et on admet sans démonstration que la suite $(\gamma_n)_{n\geqslant 1}$ est convergente ; on note γ sa limite.

a) Montrer que pour tout x réel et n assez grand, on a : $F_{G_n}(x) = \left(1 - \dfrac{1}{n}\ \mathrm{e}^{-(x+\gamma_n)}\right)^n$.

b) En déduire que pour tout x réel, on a : $\lim\limits_{n\to+\infty} F_{G_n}(x) = \mathrm{e}^{-\mathrm{e}^{-(x+\gamma)}}$.

c) Montrer que la fonction $F_G : \mathbb{R} \to \mathbb{R}$ définie par $F_G : x \mapsto \mathrm{e}^{-\mathrm{e}^{-(x+\gamma)}}$ est la fonction de répartition d'une variable aléatoire G à densité. Conclure.

7. ***a)*** Soit X une variable aléatoire à densité de fonction de répartition F_X strictement croissante. Déterminer la loi de la variable aléatoire Y définie par $Y = F_X(X)$.

b) Écrire une fonction **Scilab** d'en-tête `Gumbel` qui permet de simuler la variable aléatoire G.
On supposera que la constante γ est définie en langage **Scilab** par une constante `gamma`.
On rappelle que la fonction **Scilab** `rand()` permet de simuler la loi uniforme sur $]0,1[$.

HEC 2019 - loi de Rademacher, loi binomiale, loi uniforme, loi normale, loi de Poisson, convergence en loi, fonctions génératrice des moments, des cumulants et des probabilités

Dans ce problème, on définit et on étudie les fonctions génératrices des cumulants de variables aléatoires discrètes ou à densité.
Les cumulants d'ordre 3 et 4 permettent de définir des paramètres d'asymétrie et d'aplatissement qui viennent compléter la description usuelle d'une loi de probabilité par son espérance (paramètre de position) et sa variance (paramètre de dispersion) ; ces cumulants sont notamment utilisés pour l'évaluation des risques financiers.

Dans tout le problème :

- on note $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé et toutes les variables aléatoires introduites dans l'énoncé sont des variables aléatoires réelles définies sur $(\Omega, \mathscr{A})$;
- sous réserve d'existence, l'espérance et la variance d'une variable aléatoire X sont respectivement notées $\mathbb{E}(X)$ et $\mathbb{V}(X)$;
- pour tout variable aléatoire X et pour tout réel t pour lesquels la variable aléatoire $\mathrm{e}^{t\,X}$ admet une espérance, on pose :
$$M_X(t) = \mathbb{E}\left(\mathrm{e}^{t\,X}\right) \quad \text{et} \quad K_X(t) = \ln\big(M_X(t)\big);$$
(les fonctions M_X et K_X sont respectivement appelées la *fonction génératrice des moments* et la *fonction génératrice des cumulants* de X)
- lorsque, pour un entier $p \in \mathbb{N}^*$, la fonction K_X est de classe $\mathcal{C}^p$ sur un intervalle ouvert contenant l'origine, on appelle *cumulant d'ordre p de X*, noté $Q_p(X)$, la valeur de la dérivée $p^{\text{ème}}$ de K_X en 0 :
$$Q_p(X) \;=\; K_X^{(p)}(0).$$

Commentaire

- La fonction génératrice des moments d'une v.a.r. X est un objet classique en probabilités. Comme son nom l'indique, cette fonction permet de retrouver les moments de X (sous réserve d'existence). Plus précisément, si la v.a.r. X admet un moment d'ordre n, alors :
$$\mathbb{E}(X^n) \;=\; M_X^{(n)}(0)$$
- On ne confondra pas cette fonction avec une autre fonction classique en probabilités : la fonction génératrice (des probabilités). Cette dernière n'est définie que pour des v.a.r. X à valeurs entières et positives par la formule :
$$\forall s \in [0,1], \quad G_X(s) \;=\; \mathbb{E}(s^X) \;=\; \sum_{k=0}^{+\infty} s^k\,\mathbb{P}([X=k])$$
Cette fonction caractérise quant à elle, non pas les moments de X, mais sa loi.
Plus précisément :
$$\forall n \in \mathbb{N}, \quad \mathbb{P}([X=n]) \;=\; \frac{G_X^{(n)}(0)}{n!}$$
- L'énoncé propose de démontrer quelques propriétés de la fonction génératrice des moments pour des v.a.r. particulières. Par exemple, dans le cas d'une v.a.r. X à valeurs dans $[\![-n, n]\!]$:
  - × $\forall p \in \mathbb{N}^*$, $M_X^{(p)}(0) = \mathbb{E}(X^p)$,
  - × $\forall t \in \mathbb{R}$, $G_X(\mathrm{e}^t) = \mathrm{e}^{nt}\,M_X(t)$.
    (lien entre la fonction génératrice et la fonction génératrice des moments)

Partie I. Fonction génératrice des moments de variables aléatoires discrètes

Dans toute cette partie :

- on note n un entier supérieur ou égal à 2 ;
- toutes les variables aléatoires considérées sont discrètes à valeurs entières ;
- on note S une variable aléatoire à valeurs dans $\{-1, 1\}$ dont la loi est donnée par :

$$\mathbb{P}([S = -1]) \;=\; \mathbb{P}([S = +1]) \;=\; \frac{1}{2}.$$

1. Soit X une variable aléatoire à valeurs dans $[\![-n, n]\!]$.

a) Pour tout $t \in \mathbb{R}$, écrire $M_X(t)$ sous la forme d'une somme et en déduire que la fonction M_X est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

Démonstration.

- Soit $t \in \mathbb{R}$.
  La v.a.r. X est une v.a.r. finie. Ainsi, la v.a.r. $\mathrm{e}^{t\,X}$ est également une v.a.r. finie. Elle admet donc des moments à tout ordre, en particulier une espérance.

$$\boxed{\text{On en déduit que, pour tout } t \in \mathbb{R}, \text{ le réel } M_X(t) \text{ est bien défini.}}$$

Commentaire

On rappelle que, d'après l'énoncé : $X(\Omega) \subset [\![-n, n]\!] = \{k \mid k \in [\![-n, n]\!]\}$.
On en déduit, pour tout $t \in \mathbb{R}$:

$$\left(\mathrm{e}^{t\,X}\right)(\Omega) \;=\; \mathrm{e}^{t\,X(\Omega)} \;=\; \{\mathrm{e}^{t\,k} \mid k \in [\![-n, n]\!]\}$$

On retrouve bien que la v.a.r. $\mathrm{e}^{t\,X}$ est finie.

- Soit $t \in \mathbb{R}$. Comme X est une v.a.r. discrète, par théorème de transfert :

$$M_X(t) \;=\; \mathbb{E}\left(\mathrm{e}^{t\,X}\right) \;=\; \sum_{k=-n}^{n} \mathrm{e}^{t\,k}\,\mathbb{P}([X = k])$$

$$\boxed{\forall t \in \mathbb{R},\ M_X(t) = \sum_{k=-n}^{n} \mathrm{e}^{t\,k}\,\mathbb{P}([X = k])}$$

- La fonction M_X est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ en tant que somme, pour tout $k \in [\![-n, n]\!]$ des fonctions $t \mapsto \mathrm{e}^{t\,k}\,\mathbb{P}([X = k])$ de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

$$\boxed{\text{La fonction } M_X \text{ est de classe } \mathcal{C}^\infty \text{ sur } \mathbb{R}.}$$

Commentaire

Soit $k \in [\![-n, n]\!]$.
Il faut noter que l'expression $\mathbb{P}([X = k])$ est une constante par rapport à t. Ainsi, pour étudier la régularité de la fonction $t \mapsto \mathrm{e}^{t\,k}\,\mathbb{P}([X = k])$ on peut se contenter d'étudier la régularité de la fonction $t \mapsto \mathrm{e}^{t\,k}$ (qui, elle, est trivialement de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$).

□

b) Justifier pour tout $p \in \mathbb{N}^*$, l'égalité : $M_X^{(p)}(0) = \mathbb{E}(X^p)$.

Démonstration.

- D'après la question précédente, la fonction M_X est de classe $\mathcal{C}^\infty$, donc, pour tout $p \in \mathbb{N}^*$, la fonction $M_X^{(p)}$ existe.
- Commençons par déterminer, pour tout $p \in \mathbb{N}^*$, l'expression de $M_X^{(p)}$.
  Démontrons par récurrence : $\forall p \in \mathbb{N}^*$, $\mathcal{P}(p)$
  où $\quad \mathcal{P}(p)$: $\forall t \in \mathbb{R}$, $M_X^{(p)}(t) = \sum\limits_{k=-n}^{n} k^p \, \mathrm{e}^{k\,t} \, \mathbb{P}([X = k])$.
  - ► **Initialisation** :
    D'après la question précédente :
    $$M_X : t \mapsto \sum_{k=-n}^{n} \mathrm{e}^{k\,t} \, \mathbb{P}([X = k])$$
    Ainsi, pour tout $t \in \mathbb{R}$.
    $$M_X^{(1)}(t) \; = \; M_X'(t) \; = \; \sum_{k=-n}^{n} k \, \mathrm{e}^{k\,t} \, \mathbb{P}([X = k])$$
    D'où $\mathcal{P}(1)$.
  - ► **Hérédité** : soit $p \in \mathbb{N}^*$.
    Supposons $\mathcal{P}(p)$ et démontrons $\mathcal{P}(p{+}1)$ (*i.e.* $\forall t \in \mathbb{R}$, $M_X^{(p+1)}(t) = \sum\limits_{k=-n}^{n} k^{p+1} \, \mathrm{e}^{k\,t} \, \mathbb{P}([X = k])$).
    Par hypothèse de récurrence :
    $$M_X^{(p)} : t \mapsto \sum_{k=-n}^{n} k^p \, \mathrm{e}^{k\,t} \, \mathbb{P}([X = k])$$
    Ainsi, pour tout $t \in \mathbb{R}$:
    $$M_X^{(p+1)}(t) \; = \; \left(M_X^{(p)}\right)'(t) \; = \; \sum_{k=-n}^{n} k^p \left(k \, \mathrm{e}^{k\,t}\right) \mathbb{P}([X = k]) \; = \; \sum_{k=-n}^{n} k^{p+1} \, \mathrm{e}^{k\,t} \, \mathbb{P}([X = k])$$
    D'où $\mathcal{P}(p + 1)$.

  Par principe de récurrence : $\forall t \in \mathbb{R}$, $M_X^{(p)}(t) = \sum\limits_{k=-n}^{n} k^p \, \mathrm{e}^{k\,t} \, \mathbb{P}([X = k])$.
- Soit $p \in \mathbb{N}^*$. On en déduit :
  $$M_X^{(p)}(0) \; = \; \sum_{k=-n}^{n} k^p \, \mathrm{e}^{k \times 0} \, \mathbb{P}([X = k]) \; = \; \sum_{k=-n}^{n} k^p \, \mathbb{P}([X = k]) \; = \; \mathbb{E}(X^p)$$
  où la dernière égalité est obtenue par théorème de transfert.

$$\boxed{\text{Finalement, pour tout } p \in \mathbb{N}^* : M_X^{(p)}(0) = \mathbb{E}(X^p).}$$

Commentaire

L'obtention de l'expression de $M_X^{(p)}$ s'effectue au brouillon :

1) on commence par déterminer les premières dérivées successives de la fonction M_X :

$$\begin{array}{rcl} \forall t \in \mathbb{R}, \quad M_X'(t) & = & \sum\limits_{k=-n}^{n} k\, \mathrm{e}^{k\,t}\, \mathbb{P}([X=k]) \\ \forall t \in \mathbb{R}, \quad M_X''(t) & = & \sum\limits_{k=-n}^{n} k\left(k\, \mathrm{e}^{k\,t}\right) \mathbb{P}([X=k]) \; = \; \sum\limits_{k=-n}^{n} k^2 \mathrm{e}^{k\,t}\, \mathbb{P}([X=k]) \\ \forall t \in \mathbb{R}, \quad M_X^{(3)}(t) & = & \sum\limits_{k=-n}^{n} k^2\left(k\, \mathrm{e}^{k\,t}\right) \mathbb{P}([X=k]) \; = \; \sum\limits_{k=-n}^{n} k^3 \mathrm{e}^{k\,t}\, \mathbb{P}([X=k]) \end{array}$$

2) on en déduit une formule générale :

$$\forall t \in \mathbb{R}, \quad M_X^{(p)}(t) \; = \; \sum_{k=-n}^{n} k^p\, \mathrm{e}^{k\,t}\, \mathbb{P}([X=k])$$

□

c) Soit Y une variable aléatoire à valeurs dans $[\![-n,n]\!]$ dont la fonction génératrice des moments M_Y est la même que celle de X.
On note G_X et G_Y les deux polynômes définis par :

$$\forall x \in \mathbb{R}, \quad \left\{ \begin{array}{rcl} G_X(x) & = & \sum\limits_{k=0}^{2n} \mathbb{P}([X=k-n])\, x^k \\ G_Y(x) & = & \sum\limits_{k=0}^{2n} \mathbb{P}([Y=k-n])\, x^k \end{array} \right.$$

(i) Vérifier pour tout $t \in \mathbb{R}$, l'égalité : $G_X(\mathrm{e}^t) = \mathrm{e}^{nt}\, M_X(t)$.

Démonstration.
Soit $t \in \mathbb{R}$.

$$\begin{array}{rcll} G_X\left(\mathrm{e}^t\right) & = & \sum\limits_{k=0}^{2n} \mathbb{P}([X=k-n])\, \left(\mathrm{e}^t\right)^k & \\ & = & \sum\limits_{k=0}^{2n} \mathbb{P}([X=k-n])\, \mathrm{e}^{k\,t} & \\ & = & \sum\limits_{k=0-n}^{2n-n} \mathbb{P}([X=k])\, \mathrm{e}^{(k+n)\,t} & \textit{(par décalage d'indice)} \\ & = & \sum\limits_{k=-n}^{n} \left(\mathbb{P}([X=k])\, \mathrm{e}^{k\,t}\, \mathrm{e}^{n\,t}\right) & \\ & = & \mathrm{e}^{n\,t} \sum\limits_{k=-n}^{n} \mathbb{P}([X=k])\, \mathrm{e}^{k\,t} & \\ & = & \mathrm{e}^{n\,t}\, M_X(t) & \end{array}$$

$$\boxed{\forall t \in \mathbb{R}, G_X\left(\mathrm{e}^t\right) = \mathrm{e}^{nt}\, M_X(t)}$$

□

(ii) Justifier la relation : $\forall t \in \mathbb{R},\ G_X(\mathrm{e}^t) = G_Y(\mathrm{e}^t)$.

Démonstration.
Soit $t \in \mathbb{R}$.

$$\begin{array}{rcll} G_X\left(\mathrm{e}^t\right) & = & \mathrm{e}^{n\,t}\, M_X(t) & \textit{(d'après la question précédente)} \\ & = & \mathrm{e}^{n\,t}\, M_Y(t) & \textit{(car, d'après l'énoncé : } M_X = M_Y\textit{)} \\ & = & G_Y\left(\mathrm{e}^t\right) & \textit{(avec le même raisonnement qu'en question précédente)} \end{array}$$

$$\boxed{\forall t \in \mathbb{R},\ G_X\left(\mathrm{e}^t\right) = G_Y\left(\mathrm{e}^t\right)}$$

□

(iii) En déduire que la variable aléatoire Y suit la même loi que X.

Démonstration.

- Tout d'abord : $X(\Omega) \subset [\![-n, n]\!]$ et $Y(\Omega) \subset [\![-n, n]\!]$.
- Ensuite, d'après la question précédente :

$$\forall t \in \mathbb{R},\ G_X\left(\mathrm{e}^t\right) = G_Y\left(\mathrm{e}^t\right)$$

Ainsi, pour tout $x \in]0, +\infty[$:

$$\begin{array}{ccc} G_X\left(\mathrm{e}^{\ln(x)}\right) & = & G_Y\left(\mathrm{e}^{\ln(x)}\right) \\ \| & & \| \\ G_X(x) & & G_Y(x) \end{array}$$

- Soit $x \in]0, +\infty[$, on en déduit : $(G_X - G_Y)(x) = 0$.
  Or :

$$\begin{array}{rcl} (G_X - G_Y)(x) & = & G_X(x) - G_Y(x) \\ & = & \sum\limits_{k=0}^{2n} \mathbb{P}([X = k-n])\, x^k - \sum\limits_{k=0}^{2n} \mathbb{P}([Y = k-n])\, x^k \\ & = & \sum\limits_{k=0}^{2n} \left(\mathbb{P}([X = k-n])\, x^k - \mathbb{P}([Y = k-n])\, x^k\right) \\ & = & \sum\limits_{k=0}^{2n} \left(\mathbb{P}([X = k-n]) - \mathbb{P}([Y = k-n])\right) x^k \end{array}$$

Ainsi, le polynôme $G_X - G_Y$ admet une infinité de racines. C'est donc le polynôme nul, *i.e.* :

$$\forall k \in [\![0, 2n]\!],\ \mathbb{P}([X = k-n]) - \mathbb{P}([Y = k-n]) = 0$$

D'où :

$$\forall k \in [\![-n, n]\!],\ \mathbb{P}([X = k]) - \mathbb{P}([Y = k]) = 0$$

Ainsi :

$$\forall k \in [\![-n, n]\!],\ \mathbb{P}([X = k]) = \mathbb{P}([Y = k])$$

$$\boxed{\text{Finalement, les v.a.r. } X \text{ et } Y \text{ suivent la même loi.}}$$

□

2. Dans cette question, on note X_2 une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(2, \frac{1}{2}\right)$.
On suppose que les variables aléatoires X_2 et S sont indépendantes et on pose $Y_2 = S\, X_2$.

a) (i) Préciser l'ensemble des valeurs possibles de la variable aléatoire Y_2.

Démonstration.

- Comme $X_2 \hookrightarrow \mathcal{B}\left(2, \frac{1}{2}\right)$, on a : $X_2(\Omega) = [\![0, 2]\!] = \{0, 1, 2\}$.
- De plus : $S(\Omega) = \{-1, 1\}$.

$$\boxed{\text{On en déduit : } Y_2(\Omega) = (S\, X_2)(\Omega) \subset \{-2, -1, 0, 1, 2\} = [\![-2, 2]\!].}$$

□

(ii) Calculer les probabilités $\mathbb{P}([Y_2 = y])$ attachées aux diverses valeurs possibles y de Y_2.

Démonstration.

- Tout d'abord, on a l'égalité entre événements :

$$[Y_2 = -2] \ = \ [S\, X_2 = -2] \ = \ [S = -1] \cap [X_2 = 2]$$

On en déduit :

$$\begin{aligned} \mathbb{P}([Y_2 = -2]) &= \mathbb{P}([S = -1] \cap [X_2 = 2]) \\ &= \mathbb{P}([S = -1])\ \mathbb{P}([X_2 = 2]) \quad \textit{(car S et X_2 sont indépendantes)} \\ &= \frac{1}{2} \times \binom{2}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{2-2} \\ &= \frac{1}{2} \times 1 \times \frac{1}{4} \end{aligned}$$

$$\boxed{\mathbb{P}([Y_2 = -2]) = \frac{1}{8}}$$

- Ensuite :

$$[Y_2 = -1] \ = \ [S = -1] \cap [X_2 = 1]$$

On obtient donc :

$$\begin{aligned} \mathbb{P}([Y_2 = -1]) &= \mathbb{P}([S = -1] \cap [X_2 = 1]) \\ &= \mathbb{P}([S = -1])\ \mathbb{P}([X_2 = 1]) \quad \textit{(car S et X_2 sont indépendantes)} \\ &= \frac{1}{2} \times \binom{2}{1} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{2-1} \\ &= \frac{1}{2} \times 2 \times \frac{1}{4} \end{aligned}$$

$$\boxed{\mathbb{P}([Y_2 = -1]) = \frac{1}{4}}$$

- De même :

$$[Y_2 = 1] = [S = 1] \cap [X_2 = 1] \quad \text{et} \quad [Y_2 = 2] = [S = 1] \cap [X_2 = 2]$$

$$\boxed{\text{Ainsi : } \mathbb{P}([Y_2 = 1]) = \frac{1}{4} \quad \text{et} \quad \mathbb{P}([Y_2 = 2]) = \frac{1}{8}.}$$

- Enfin : $[Y_2 = 0] = [X_2 = 0]$. Ainsi :

$$\mathbb{P}([Y_2 = 0]) \; = \; \mathbb{P}([X_2 = 0]) \; = \; \binom{2}{0} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^2 \; = \; \frac{1}{4}$$

$$\boxed{\mathbb{P}([Y_2 = 0]) = \frac{1}{4}}$$

Commentaire

On pouvait aussi remarquer que la famille $([Y_2 = k])_{k \in [\![-2,2]\!]}$ est un système complet d'événements. D'où :

$$\begin{aligned} \mathbb{P}([Y_2 = 0]) \; &= \; 1 - \Big(\mathbb{P}([Y_2 = -2]) + \mathbb{P}([Y_2 = -1]) + \mathbb{P}([Y_2 = 1]) + \mathbb{P}([Y_2 = 2])\Big) \\ &= \; 1 - \left(\frac{1}{8} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8}\right) \; = \; \frac{1}{4} \end{aligned}$$

□

b) Vérifier que la variable aléatoire $X_2 - (S + 1)$ suit la même loi que Y_2.

Démonstration.

- Tout d'abord, comme $S(\Omega) = \{-1, 1\}$, on a : $(S + 1)(\Omega) = \{0, 2\}$.
  De plus $X_2(\Omega) = \{0, 1, 2\}$. D'où, en notant $T_2 = X_2 - (S + 1)$: $T_2(\Omega) \subset \{-2, -1, 0, 1, 2\}$.

  $$\boxed{T_2(\Omega) \subset [\![-2, 2]\!] \text{ (on rappelle : } Y_2(\Omega) \subset [\![-2, 2]\!] \text{ d'après } \textbf{\textit{2.a)(i)}}\text{)}}$$

- Soit $k \in [\![-2, 2]\!]$.
  La famille $([S = -1], [S = 1])$ forme un système complet d'événements.
  Ainsi, par formule des probabilités totales :

$$\begin{aligned} & \mathbb{P}([T_2 = k]) \\ = \; & \mathbb{P}([T_2 = k] \cap [S = -1]) + \mathbb{P}([T_2 = k] \cap [S = 1]) \\ = \; & \mathbb{P}([X_2 - (S + 1) = k] \cap [S = -1]) + \mathbb{P}([X_2 - (S + 1) = k] \cap [S = 1]) \\ = \; & \mathbb{P}([X_2 - (-1 + 1) = k] \cap [S = -1]) + \mathbb{P}([X_2 - (1 + 1) = k] \cap [S = 1]) \\ = \; & \mathbb{P}([X_2 = k] \cap [S = -1]) + \mathbb{P}([X_2 = k + 2] \cap [S = 1]) \\ = \; & \mathbb{P}([X_2 = k]) \; \mathbb{P}([S = -1]) + \mathbb{P}([X_2 = k + 2]) \; \mathbb{P}([S = 1]) && \textit{(car } X_2 \textit{ et } S \textit{ sont indépendantes)} \\ = \; & \frac{1}{2} \; \mathbb{P}([X_2 = k]) + \frac{1}{2} \; \mathbb{P}([X_2 = k + 2]) \end{aligned}$$

- On obtient ainsi :

$$\begin{aligned}
\mathbb{P}([T_2=-2]) &= \frac{1}{2}\,\mathbb{P}([X_2=-2]) + \frac{1}{2}\,\mathbb{P}([X_2=0]) = \cancel{\frac{1}{2}\,\mathbb{P}(\varnothing)} + \frac{1}{2}\times\frac{1}{4} = \frac{1}{8} = \mathbb{P}([Y_2=-2])\\
\mathbb{P}([T_2=-1]) &= \frac{1}{2}\,\mathbb{P}([X_2=-1]) + \frac{1}{2}\,\mathbb{P}([X_2=1]) = \cancel{\frac{1}{2}\,\mathbb{P}(\varnothing)} + \frac{1}{2}\times\frac{1}{2} = \frac{1}{4} = \mathbb{P}([Y_2=-1])\\
\mathbb{P}([T_2=0]) &= \frac{1}{2}\,\mathbb{P}([X_2=0]) + \frac{1}{2}\,\mathbb{P}([X_2=2]) = \frac{1}{2}\times\frac{1}{4} + \frac{1}{2}\times\frac{1}{4} = \frac{1}{4} = \mathbb{P}([Y_2=0])\\
\mathbb{P}([T_2=1]) &= \frac{1}{2}\,\mathbb{P}([X_2=1]) + \frac{1}{2}\,\mathbb{P}([X_2=3]) = \frac{1}{2}\times\frac{1}{2} + \cancel{\frac{1}{2}\,\mathbb{P}(\varnothing)} = \frac{1}{4} = \mathbb{P}([Y_2=1])\\
\mathbb{P}([T_2=2]) &= \frac{1}{2}\,\mathbb{P}([X_2=2]) + \frac{1}{2}\,\mathbb{P}([X_2=4]) = \frac{1}{2}\times\frac{1}{4} + \cancel{\frac{1}{2}\,\mathbb{P}(\varnothing)} = \frac{1}{8} = \mathbb{P}([Y_2=2])
\end{aligned}$$

$$\boxed{\text{On en déduit que } T_2 = X_2 - (S+1) \text{ et } Y_2 \text{ suivent la même loi.}}$$

Commentaire

- Lors d'une rédaction classique de détermination de la loi d'une somme, après obtention de la relation :

$$\mathbb{P}([T_2=k]) = \frac{1}{2}\,\mathbb{P}([X_2=k]) + \frac{1}{2}\,\mathbb{P}([X_2=k+2])$$

  On cherche à déterminer si les événements $[X_2=k]$ et/ou $[X_2=k+2]$ sont l'événement impossible.
- Ici, cela varie suivant les valeurs de k. C'est pour cela qu'on effectue le calcul direct des probabilités $\mathbb{P}([T_2=k])$ après l'obtention de la formule précédente.

□

3. Le script **Scilab** suivant permet d'effectuer des simulations de la variable aléatoire Y_2 définie dans la question précédente.

```
n = 10
X = grand(n,2,'bin',2,0.5)
B = grand(n,2,'bin',1,0.5)
S = 2 * B - ones(n,2)
Z1 = [S(1:n,1) .* X(1:n,1) , X(1:n,1) - S(1:n,1) - ones(n,1)]
Z2 = [S(1:n,1) .* X(1:n,1) , X(1:n,2) - S(1:n,2) - ones(n,1)]
```

a) Que contiennent les variables `X` et `S` après l'exécution des quatre premières instructions ?

Démonstration.

- L'instruction `X = grand(n,2,'bin',2,0.5)` permet de stocker dans la variable `X` une matrice à `n` lignes et 2 colonnes où chaque colonne contient une observation d'un `n`-échantillon de loi $\mathcal{B}\left(2,\frac{1}{2}\right)$, c'est-à-dire un `n`-échantillon de la v.a.r. X. Plus précisément, la première colonne de `X` contient une observation $(c_1,\dots,c_n)$ d'un `n`-échantillon $(C_1,\dots,C_n)$ de X, et la deuxième colonne de `X` contient une observation $(c'_1,\dots,c'_n)$ d'un `n`-échantillon $(c'_1,\dots,C'_n)$ de `X`.
(les v.a.r. C_i et C'_i sont indépendantes et ont même loi que la v.a.r. X)

- L'instruction `B = grand(n,2,'bin',1,0.5)` permet de stocker dans la variable `B` une matrice à `n` lignes et 2 colonnes où la première colonne contient une observation $(b_1, \dots, b_n)$ d'un `n`-échantillon $(B_1, \dots, B_n)$ de loi $\mathcal{B}\left(\frac{1}{2}\right)$, et la deuxième colonne contient une observation $(b'_1, \dots, b'_n)$ d'un `n`-échantillon $(B'_1, \dots, B'_n)$ de loi $\mathcal{B}\left(\frac{1}{2}\right)$.
  (les v.a.r. B_i et B'_i sont indépendantes et ont même loi $\mathcal{B}\left(\frac{1}{2}\right)$)
- On commence par rappeler que l'instruction `ones(n,2)` permet d'obtenir une matrice à `n` lignes et 2 colonnes dont tous les coefficients sont égaux à 1.
  On en déduit que l'instruction `S = 2 * B - ones(n,2)` permet de stocker dans la variable `S` une matrice à `n` lignes et 2 colonnes.
  La première colonne contient l'observation $(s_1, \dots, s_n) = (2\,b_1 - 1, \dots, 2\,b_n - 1)$ du `n`-échantillon $(2\,B_1 - 1, \dots, 2\,B_n - 1)$.
  La deuxième colonne contient l'observation $(s'_1, \dots, s'_n) = (2\,b'_1 - 1, \dots, 2\,b'_n - 1)$ du `n`-échantillon $(2\,B'_1 - 1, \dots, 2\,B'_n - 1)$.
- Les v.a.r. $2\,B_i - 1$ et $2\,B'_i - 1$ sont indépendantes et de même loi. Déterminons cette loi : on note B une v.a.r. de loi $\mathcal{B}\left(\frac{1}{2}\right)$ et on cherche la loi de $V = 2\,B - 1$.

  × Tout d'abord, comme $B \hookrightarrow \mathcal{B}\left(\frac{1}{2}\right)$, on a : $B(\Omega) = \{0, 1\}$.
  On en déduit : $V(\Omega) = \{2 \times 0 - 1, 2 \times 1 - 1\} = \{-1, 1\}$.

  × De plus :
  $$[V = 1] \; = \; [2\,B - 1 = 1] \; = \; [2\,B = 2] \; = \; [B = 1]$$
  D'où : $\mathbb{P}([V = 1]) = \mathbb{P}([B = 1]) = \frac{1}{2}$

  × Enfin, comme la famille $([V = -1], [V = 1])$ est un système complet d'événements :
  $$\mathbb{P}([V = -1]) \; = \; 1 - \mathbb{P}([V = 1]) \; = \; 1 - \frac{1}{2} \; = \; \frac{1}{2}$$

  On en déduit que la v.a.r. $V = 2\,B - 1$ suit la même loi que la v.a.r. S.
- La variable `S` contient donc une matrice à `n` lignes et 2 colonnes, où la première colonne contient une observation $(s_1, \dots, s_n)$ d'un `n`-échantillon $(S_1, \dots, S_n)$ de la v.a.r. S, et la deuxième colonne contient une observation $(s'_1, \dots, s'_n)$ du `n`-échantillon $(S'_1, \dots, S'_n)$ de la v.a.r. S.
  (les v.a.r. S_i et S'_i sont indépendantes et ont même loi que la v.a.r. S)

Commentaire

- Comme souvent dans les sujets HEC, l'une des difficultés provient de la prise d'initiatives nécessaires pour répondre à une question.
- Dans cette question par exemple, déterminer la loi de la v.a.r. $V = 2\,B - 1$ n'est pas difficile. C'est prendre l'initiative de déterminer cette loi qui constitue la difficulté.

□

b) Expliquer pourquoi, après l'exécution des six instructions, chacun des coefficients des matrices `Z1` et `Z2` contient une simulation de la variable aléatoire Y_2.

Démonstration.

- L'instruction `Z1 = [S(1:n,1) .* X(1:n,1) , X(1:n,1) - S(1:n,1) - ones(n,1)]` permet de stocker dans la variable `Z1` une matrice à `n` lignes et 2 colonnes.
  Le contenu de la première colonne est donné par la commande `S(1:n,1) .* X(1:n,1)` et celui de la deuxième colonne par `X(1:n,1) - S(1:n,1) - ones(n,1)`.
- L'instruction `S(1:n,1) .* X(1:n,1)` permet d'obtenir une matrice colonne à `n` lignes contenant l'observation $(d_1, \dots, d_n) = (s_1 \times c_1, \dots, s_n \times c_n)$ du `n`-échantillon $(S_1\, C_1, \dots, S_n\, C_n)$.
  Or, les S_i suivent la même loi que S et les C_i suivent la même loi que X_2. Ainsi, les $S_i\, C_i$ suivent la même loi que Y_2.
  La première colonne de `Z1` contient donc une observation $(d_1, \dots, d_n)$ d'un `n`-échantillon $(D_1, \dots, D_n)$ de Y_2.
  (les v.a.r. D_i sont indépendantes et ont même loi que la v.a.r. Y_2)
- L'instruction `X(1:n,1) - S(1:n,1) - ones(n,1)` permet d'obtenir une matrice colonne à `n` lignes contenant l'observation $(d'_1, \dots, d'_n) = (c_1 - s_1 - 1, \dots, c_n - s_n - 1)$ du `n`-échantillon $(C_1 - S_1 - 1, \dots, C_n - S_n - 1)$.
  Or, les S_i suivent la même loi que S et les C_i suivent la même loi que X_2. Ainsi, les $C_i - S_i - 1$ suivent la même loi que $X_2 - S - 1$. De plus, d'après la question ***2.b)***, la v.a.r. $X_2 - S - 1$ suit la même loi que Y_2.
  La deuxième colonne de `Z1` contient donc une observation $(d'_1, \dots, d'_n)$ d'un `n`-échantillon $(D'_1, \dots, D'_n)$ de Y_2.
  (les v.a.r. D'_i sont indépendantes et ont même loi que la v.a.r. Y_2)
- De même, l'instruction `Z2 = [S(1:n,1) .* X(1:n,1) , X(1:n,2) - S(1:n,2) - ones(n,1)]` permet de stocker dans la variable `Z2` une matrice à `n` lignes et 2 colonnes.
  Le contenu de la première colonne est donné par la commande `S(1:n,1) .* X(1:n,1)` et celui de la deuxième colonne par `X(1:n,2) - S(1:n,2) - ones(n,1)`.
- La première colonne de `Z2` est identique à celle de `Z1`, donc la première colonne de `Z2` contient l'observation $(d_1, \dots, d_n)$ du `n`-échantillon $(D_1, \dots, D_n)$ de Y_2.
- L'instruction `X(1:n,2) - S(1:n,2) - ones(n,1)` permet d'obtenir une matrice colonne à `n` lignes contenant l'observation $(d''_1, \dots, d''_n) = (c'_1 - s'_1 - 1, \dots, c'_n - s'_n - 1)$ du `n`-échantillon $(C'_1 - S'_1 - 1, \dots, C'_n - S'_n - 1)$.
  Or, les S'_i suivent la même loi que S et les C'_i suivent la même loi que X_2. Ainsi, les $C'_i - S'_i - 1$ suivent la même loi que $X_2 - S - 1$, donc que Y_2.
  La deuxième colonne de `Z2` contient donc une observation $(d''_1, \dots, d''_n)$ d'un `n`-échantillon $(D''_1, \dots, D''_n)$ de Y_2.
  (les v.a.r. D''_i sont indépendantes et ont même loi que la v.a.r. Y_2)

Finalement, chacun des coefficients des matrices `Z1` et `Z2` contient une simulation de la v.a.r. Y_2.

Commentaire

On aurait pu utiliser davantage les commandes Scilab. En effet, l'appel classique permettant d'extraire la première colonne de la matrice `S` est plutôt `S(:,1)` (que `S(1:n,1)`). $\square$

c) On modifie la première ligne du script précédent en affectant à `n` une valeur beaucoup plus grande que 10 (par exemple, 100000) et en lui adjoignant les deux instructions 7 et 8 suivantes :

```
7  p1 = length(find(Z1(1:n,1) == Z1(1:n,2))) / n
8  p2 = length(find(Z2(1:n,1) == Z2(1:n,2))) / n
```

Quelles valeurs numériques approchées la loi faible des grands nombres permet-elle de fournir pour `p1` et `p2` après l'exécution des huit lignes du nouveau script ?

Dans le langage **Scilab**, la fonction `length` fournit la « longueur » d'un vecteur ou d'une matrice et la fonction `find` calcule les positions des coefficients d'une matrice pour lesquels une propriété est vraie, comme l'illustre le script suivant :

```
--> A = [1 ; 2 ; 0 ; 4]
--> B = [2 ; 2 ; 4 ; 3]
--> length(A)
   ans = 4.
--> length([A , B])
   ans = 8.
--> find(A < B)
   ans = 1. 3. // car 1 < 2 et 0 < 4, alors que 2 ⩾ 2 et 4 ⩾ 3
```

Démonstration.

- Commençons par commenter l'instruction :

```
7  p1 = length(find(Z1(1:n,1) == Z1(1:n,2))) / n
```

  × L'instruction `find(Z1(1:n,1) == Z1(1:n,2))` permet d'obtenir une matrice ligne contenant les positions des coefficients des matrices `Z1(1:n,1)` et `Z1(1:n,2)` égaux. Autrement dit, on obtient une matrice ligne contenant les indices i tels que $d_i = d'_i$.
  (on rappelle que $(d_1, \dots, d_n)$ est une observation d'un `n`*-échantillon $(D_1, \dots, D_n)$ de $S\,X_2$ et $(d'_1, \dots, d'_n)$ est une observation d'un* `n`*-échantillon $(D'_1, \dots, D'_n)$ de $X_2 - S - 1$)*

  × L'instruction `length(find(Z1(1:n,1) == Z1(1:n,2)))` permet d'obtenir la longueur de la matrice précédente. Ainsi, on obtient le nombre de fois où $d_i = d'_i$, pour $i \in [\![1, \mathtt{n}]\!]$.

  × Enfin, on divise ce nombre par la taille `n` de l'observation.
  Or, par loi faible des grands nombres (LfGN) :

$$\frac{\text{nombre de fois où } d_i = d'_i}{\text{taille de l'observation}} \simeq \mathbb{P}([S\,X_2 = X_2 - S - 1])$$

$$\boxed{\text{La variable } \mathtt{p1} \text{ contient une valeur approchée de } \mathbb{P}([S\,X_2 = X_2 - S - 1]).}$$

- Commentons ensuite l'instruction :

```
8  p2 = length(find(Z2(1:n,1) == Z2(1:n,2))) / n
```

  × L'instruction `find(Z2(1:n,1) == Z2(1:n,2))` permet d'obtenir une matrice ligne contenant les positions des coefficients des matrices `Z1(2:n,1)` et `Z1(2:n,2)` égaux. Autrement dit, on obtient une matrice ligne contenant les indices i tels que $d_i = d''_i$.
  (on rappelle que $(d_1, \dots, d_n)$ est une observation d'un `n`*-échantillon $(D_1, \dots, D_n)$ de $S\,X_2$ et $(d''_1, \dots, d''_n)$ est une observation d'un* `n`*-échantillon $(D''_1, \dots, D''_n)$ de $X'_2 - S - 1$, où la v.a.r. X'_2 suit la même loi que X_2)*

  × L'instruction `length(find(Z2(1:n,1) == Z2(1:n,2)))` permet d'obtenir la longueur de la matrice précédente. Ainsi, on obtient le nombre de fois où $d_i = d''_i$, pour $i \in [\![1, \mathtt{n}]\!]$.

× Enfin, on divise ce nombre par la taille `n` de l'observation.
Or, par loi faible des grands nombres (LfGN) :

$$\frac{\text{nombre de fois où } d_i = d_i''}{\text{taille de l'observation}} \simeq \mathbb{P}\big(\big[S\,X_2 = X_2' - S - 1\big]\big)$$

$$\boxed{\begin{array}{c}\text{La variable } \texttt{p2} \text{ contient une valeur approchée de } \mathbb{P}([S\,X_2 = X_2' - S - 1]) \\ \text{où la v.a.r. } X_2' \text{ suit la même loi que } X_2.\end{array}}$$

Commentaire

- Le programme proposé par l'énoncé n'est ici rien d'autre qu'une illustration de l'idée naturelle pour obtenir une approximation de $\mathbb{P}([S\,X_2 = X_2 - S - 1])$:
  - × simuler un grand nombre de fois ($\texttt{n} = 100000$) les v.a.r. $S\,X_2$ et $X_2 - S - 1$.
    Formellement, on souhaite obtenir une observation $(d_1, \dots, d_n)$ d'un `n`-échantillon $(D_1, \dots, D_n)$ de la v.a.r. $S\,X_2$, et une observation $(d_1', \dots, d_n')$ d'un `n`-échantillon $(D_1', \dots, D_n')$ de la v.a.r. $X_2 - S - 1$.
  - × de compter le nombre de fois où $d_i = d_i'$, pour $i \in [\![1, \texttt{n}]\!]$.
- L'objectif de cette question **Scilab** est de revenir sur un point important en probabilités :

$$X \text{ et } Y \text{ ont même loi} \quad \not\Rightarrow \quad X = Y$$

  (mais bien sûr, si $X = Y$, alors X et Y ont même loi)
- On peut avoir la confirmation du point précédent en exécutant le programme **Scilab**.
  On obtient :

```
p1 =
    0.1261
p2 =
    0.2179
```

  On constate qu'on a $\texttt{p1} \neq 1$ et $\texttt{p2} \neq 1$. Ainsi, ici :
  - × les v.a.r. $S\,X_2$ et $X_2 - S - 1$ ont même loi,
  - × mais $\mathbb{P}([S\,X_2 = X_2 - S - 1]) \neq 1$. Cela démontre en particulier : $S\,X_2 \neq X_2 - S - 1$. □

4. Dans cette question, on note X_n une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(n, \frac{1}{2}\right)$.
On suppose que les variables aléatoires X_n et S sont indépendantes et on pose $Y_n = S\,X_n$.

a) Justifier que la fonction M_{X_n} est définie sur $\mathbb{R}$ et calculer $M_{X_n}(t)$ pour tout $t \in \mathbb{R}$.

Démonstration.

- Soit $t \in \mathbb{R}$.
  La v.a.r. X_n est une v.a.r. finie. Ainsi, la v.a.r. $e^{t\,X_n}$ est également une v.a.r. finie. Elle admet donc des moments à tout ordre, en particulier une espérance.

$$\boxed{\text{On en déduit que la fonction } M_{X_n} \text{ défini sur } \mathbb{R}.}$$

- Soit $t \in \mathbb{R}$.

$$\begin{array}{rcll} M_{X_n}(t) & = & \mathbb{E}\left(\mathrm{e}^{t\,X_n}\right) & \\ & = & \sum\limits_{k=0}^{n} \mathrm{e}^{k\,t}\ \mathbb{P}([X_n = k]) & \textit{(par théorème de transfert } (X_n(\Omega) = [\![0,n]\!])\textit{)} \\ & = & \sum\limits_{k=0}^{n} \mathrm{e}^{k\,t} \binom{n}{k} \left(\frac{1}{2}\right)^{k} \left(\frac{1}{2}\right)^{n-k} & \textit{(car } X_n \hookrightarrow \mathcal{B}\left(n, \frac{1}{2}\right)\textit{)} \\ & = & \sum\limits_{k=0}^{n} \binom{n}{k} \mathrm{e}^{k\,t} \left(\frac{1}{2}\right)^{n} & \\ & = & \frac{1}{2^n} \sum\limits_{k=0}^{n} \binom{n}{k} \left(\mathrm{e}^{t}\right)^{k}\ 1^{n-k} & \\ & = & \frac{1}{2^n} \left(\mathrm{e}^{t} + 1\right)^{n} & \textit{(d'après la formule du binôme de Newton)} \end{array}$$

$$\boxed{\forall t \in \mathbb{R},\ M_{X_n}(t) = \frac{1}{2^n}\left(1 + \mathrm{e}^{t}\right)^{n}}$$

□

b) Montrer que la fonction M_{Y_n} est donnée par : $\forall t \in \mathbb{R},\ M_{Y_n}(t) = \dfrac{1}{2^{n+1}}\ \big((1+\mathrm{e}^{t})^n + (1+\mathrm{e}^{-t})^n\big)$.

Démonstration.

- Soit $t \in \mathbb{R}$.
  La v.a.r. Y_n est une v.a.r. finie (car les v.a.r. X_n et S le sont). Ainsi, la v.a.r. $\mathrm{e}^{t\,Y_n}$ est également une v.a.r. finie. Elle admet donc des moments à tout ordre, en particulier une espérance.

$$\boxed{\text{On en déduit que la fonction } M_{Y_n} \text{ défini sur } \mathbb{R}.}$$

- Déterminons la loi de Y_n.
  × Tout d'abord, comme $X_n(\Omega) = [\![0, n]\!]$ et $S(\Omega) = \{-1, 1\}$, on a : $Y_n(\Omega) \subset [\![-n, n]\!]$.
  × Soit $k \in [\![-n, n]\!]$.
  La famille $([S = -1], [S = 1])$ forme un système complet d'événements.
  Ainsi, par formule des probabilités totales :

$$\begin{array}{rcll} \mathbb{P}([Y_n = k]) & = & \mathbb{P}([S=-1] \cap [Y_n = k]) + \mathbb{P}([S=1] \cap [Y_n = k]) & \\ & = & \mathbb{P}([S=-1] \cap [S\,X_n = k]) + \mathbb{P}([S=1] \cap [S\,X_n = k]) & \\ & = & \mathbb{P}([S=-1] \cap [-X_n = k]) + \mathbb{P}([S=1] \cap [X_n = k]) & \\ & = & \mathbb{P}([S=-1])\ \mathbb{P}([X_n = -k]) + \mathbb{P}([S=1])\ \mathbb{P}([X_n = k]) & \textit{(car les v.a.r. } S \textit{ et } X_n \textit{ sont indépendantes)} \\ & = & \frac{1}{2}\ \mathbb{P}([X_n = -k]) + \frac{1}{2}\ \mathbb{P}([X_n = k]) & \end{array}$$

  Trois cas se présentent alors :

  - si $k \in [\![-n, 0[\![$, alors $[X_n = k] = \varnothing$. Donc :

$$\mathbb{P}([Y_n = k]) \ = \ \frac{1}{2}\ \mathbb{P}([X_n = -k]) + \cancel{\frac{1}{2}\ \mathbb{P}(\varnothing)} \ = \ \frac{1}{2}\ \mathbb{P}([X_n = -k])$$

- si $k = 0$, alors :

$$\begin{aligned}\mathbb{P}([Y_n = 0]) &= \frac{1}{2}\,\mathbb{P}([X_n = 0]) + \frac{1}{2}\,\mathbb{P}([X_n = 0]) \\ &= \mathbb{P}([X_n = 0]) \\ &= \binom{n}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{n-0} \\ &= \frac{1}{2^n}\end{aligned}$$

- si $k \in]\!]0, n]\!]$, alors $[X_n = -k] = \varnothing$. Donc :

$$\mathbb{P}([Y_n = k]) = \cancel{\frac{1}{2}\,\mathbb{P}(\varnothing)} + \frac{1}{2}\,\mathbb{P}([X_n = k]) = \frac{1}{2}\,\mathbb{P}([X_n = k])$$

- Soit $t \in \mathbb{R}$.

$$\begin{aligned}M_{Y_n}(t) &= \mathbb{E}\left(\mathrm{e}^{t\,Y_n}\right) \\ &= \sum_{k=-n}^{n} \mathrm{e}^{k\,t}\,\mathbb{P}([Y_n = k]) \\ &= \sum_{k=-n}^{-1}\left(\mathrm{e}^{k\,t}\,\mathbb{P}([Y_n = k])\right) + \mathrm{e}^{0\,t}\,\mathbb{P}([Y_n = 0]) + \sum_{k=1}^{n}\left(\mathrm{e}^{k\,t}\,\mathbb{P}([Y_n = k])\right) \\ &= \frac{1}{2}\sum_{k=-n}^{-1}\left(\mathrm{e}^{k\,t}\,\mathbb{P}([X_n = -k])\right) + \frac{1}{2^n} + \frac{1}{2}\sum_{k=1}^{n}\left(\mathrm{e}^{k\,t}\,\mathbb{P}([X_n = k])\right)\end{aligned}$$

× Tout d'abord :

$$\begin{aligned}\sum_{k=-n}^{-1} \mathrm{e}^{k\,t}\,\mathbb{P}([X_n = -k]) &= \sum_{j=1}^{n} \mathrm{e}^{-j\,t}\,\mathbb{P}([X_n = j]) && \textit{(par changement d'indice } j = -k\textit{)} \\ &= \sum_{j=1}^{n} \mathrm{e}^{-j\,t}\binom{n}{j}\left(\frac{1}{2}\right)^j\left(\frac{1}{2}\right)^{n-j} \\ &= \frac{1}{2^n}\sum_{j=1}^{n}\binom{n}{j}\left(\mathrm{e}^{-t}\right)^j\,1^{n-j} \\ &= \frac{1}{2^n}\left(\sum_{j=0}^{n}\left(\binom{n}{j}\left(\mathrm{e}^{-t}\right)^j\,1^{n-j}\right) - 1\right) \\ &= \frac{1}{2^n}\left(\left(\mathrm{e}^{-t}+1\right)^n - 1\right) && \textit{(d'après la formule du binôme de Newton)}\end{aligned}$$

× De même :

$$\begin{aligned}\sum_{k=1}^{n} \mathrm{e}^{k\,t}\,\mathbb{P}([X_n = k]) &= \sum_{k=1}^{n} \mathrm{e}^{k\,t}\binom{n}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k} \\ &= \frac{1}{2^n}\sum_{k=1}^{n}\binom{n}{k}\left(\mathrm{e}^{t}\right)^k\,1^{n-k} \\ &= \frac{1}{2^n}\left(\sum_{k=0}^{n}\left(\binom{n}{k}\left(\mathrm{e}^{t}\right)^k\,1^{n-k}\right) - 1\right) \\ &= \frac{1}{2^n}\left(\left(\mathrm{e}^{t}+1\right)^n - 1\right) && \textit{(d'après la formule du binôme de Newton)}\end{aligned}$$

- On obtient alors :

$$\begin{aligned}
M_{Y_n}(t) &= \frac{1}{2} \times \frac{1}{2^n}\left((1+\mathrm{e}^{-t})^n - 1\right) + \frac{1}{2^n} + \frac{1}{2} \times \frac{1}{2^n}\left((1+\mathrm{e}^{t})^n - 1\right) \\
&= \frac{1}{2^{n+1}}(1+\mathrm{e}^{-t})^n - \frac{1}{2^{n+1}} + \frac{1}{2^n} + \frac{1}{2^{n+1}}(1+\mathrm{e}^{t})^n - \frac{1}{2^{n+1}} \\
&= \frac{1}{2^{n+1}}\left((1+\mathrm{e}^{-t})^n + (1+\mathrm{e}^{t})^n\right) - \frac{\cancel{2}}{2^{n+\cancel{1}}} + \frac{1}{2^n} \\
&= \frac{1}{2^{n+1}}\left((1+\mathrm{e}^{-t})^n + (1+\mathrm{e}^{t})^n\right) - \cancel{\frac{1}{2^n}} + \cancel{\frac{1}{2^n}}
\end{aligned}$$

$$\boxed{\forall t \in \mathbb{R},\ M_{Y_n}(t) = \frac{1}{2^{n+1}}\ \left((1+\mathrm{e}^{t})^n + (1+\mathrm{e}^{-t})^n\right)}$$

□

c) En utilisant l'égalité $(1+\mathrm{e}^{-t})^n = \mathrm{e}^{-nt}\,(1+\mathrm{e}^{t})^n$, montrer que Y_n suit la même loi que la différence $X_n - H_n$, où H_n est une variable aléatoire indépendante de X_n dont on précisera la loi.

Démonstration.

- Soit $t \in \mathbb{R}$. Démontrons l'égalité : $(1+\mathrm{e}^{-t})^n = \mathrm{e}^{-nt}(1+\mathrm{e}^{t})^n$.

$$\begin{aligned}
\mathrm{e}^{-n\,t}\ (1+\mathrm{e}^{t})^n &= \left(\mathrm{e}^{-t}\right)^n\ (1+\mathrm{e}^{t})^n \\
&= \left(\mathrm{e}^{-t}(1+\mathrm{e}^{t})\right)^n \\
&= (\mathrm{e}^{-t}+1)^n
\end{aligned}$$

$$\boxed{\forall t \in \mathbb{R},\ (1+\mathrm{e}^{-t})^n = \mathrm{e}^{-nt}(1+\mathrm{e}^{t})^n}$$

- Il s'agit de démontrer que les v.a.r. Y_n et $X_n - H_n$ ont même loi. On peut déjà remarquer : $Y(\Omega) \subset [\![-n,n]\!]$. Or, en question ***1.c)***, on a démontré que, si deux v.a.r. X et Y vérifient :
  × $X(\Omega) \subset [\![-n,n]\!]$,
  × $Y(\Omega) \subset [\![-n,n]\!]$,
  × $M_X \ = \ M_Y$,

  alors X et Y ont même loi.
  Pour répondre à cette question, il suffit donc de trouver une v.a.r. H_n indépendante de X_n telle que :
  $$(X_n - H_n)(\Omega) \subset [\![-n,n]\!] \quad \text{et} \quad M_{Y_n} = M_{X_n - H_n}$$
  Pour déterminer une telle v.a.r. H_n, on commencera par supposer qu'elle existe pour en déduire des propriétés sur sa loi, puis on choisira une v.a.r. vérifiant ces propriétés, et enfin on établira que l'on est bien dans le cadre d'application de la question ***1.c)***.
- Supposons qu'il existe une telle v.a.r. H_n. Alors, soit $t \in \mathbb{R}$:
  × d'une part :

$$\begin{aligned}
M_{X_n - H_n}(t) &= \mathbb{E}\left(\mathrm{e}^{t\,(X_n - H_n)}\right) \\
&= \mathbb{E}\left(\mathrm{e}^{t\,X_n}\ \mathrm{e}^{-t\,H_n}\right) \\
&= \mathbb{E}\left(\mathrm{e}^{t\,X_n}\right)\ \mathbb{E}\left(\mathrm{e}^{-t\,H_n}\right) && \textit{(car, par lemme des coalitions, les v.a.r. } \mathrm{e}^{t\,X_n} \textit{ et } \mathrm{e}^{-t\,H_n} \textit{ sont indépendantes)} \\
&= M_{X_n}(t)\ \mathbb{E}\left(\mathrm{e}^{-t\,H_n}\right) \\
&= \frac{1}{2^n}\ (1+\mathrm{e}^{t})^n\ \mathbb{E}\left(\mathrm{e}^{-t\,H_n}\right) && \textit{(d'après la question 4.a))}
\end{aligned}$$

× d'autre part :

$$\begin{aligned} M_{Y_n}(t) &= \frac{1}{2^{n+1}} \left((1+\mathrm{e}^t)^n + (1+\mathrm{e}^{-t})^n\right) && \textit{(d'après la question précédente)} \\ &= \frac{1}{2^{n+1}} \left((1+\mathrm{e}^t)^n + \mathrm{e}^{-nt}(1+\mathrm{e}^t)^n\right) && \textit{(d'après l'indication de l'énoncé)} \\ &= \frac{1}{2^{n+1}} \, (1+\mathrm{e}^t)^n \left(1+\mathrm{e}^{-nt}\right) \\ &= \frac{1}{2^n} \, (1+\mathrm{e}^t)^n \times \frac{1}{2} \, \left(1+\mathrm{e}^{-nt}\right) \end{aligned}$$

Ainsi, si H_n vérifie les propriétés souhaitées :

$$\begin{aligned} M_{X_n-H_n}(t) = M_{Y_n}(t) \quad &\Leftrightarrow \quad \cancel{\frac{1}{2^n} \, (1+\mathrm{e}^t)^n} \, \mathbb{E}\left(\mathrm{e}^{-t\,H_n}\right) = \cancel{\frac{1}{2^n} \, (1+\mathrm{e}^t)^n} \times \frac{1}{2} \, \left(1+\mathrm{e}^{-nt}\right) \\ &\Leftrightarrow \quad \mathbb{E}\left(\mathrm{e}^{-t\,H_n}\right) = \frac{1}{2} \left(1+\mathrm{e}^{-nt}\right) \\ &\Leftrightarrow \quad \mathbb{E}\left(\mathrm{e}^{-t\,H_n}\right) = \mathrm{e}^{-0\,t} \, \frac{1}{2} + \mathrm{e}^{-nt} \, \frac{1}{2} \end{aligned}$$

- On semble faire apparaître ici le théorème de transfert. On choisit donc une v.a.r. H_n qui vérifie cette égalité. On choisit alors une v.a.r. H_n, indépendante de X_n telle que :
  × $H_n(\Omega) = \{0, n\}$,
  × $\mathbb{P}([H_n = 0]) = \mathbb{P}([H_n = n]) = \dfrac{1}{2}$.
- Vérifions qu'avec cette v.a.r. H_n, nous sommes dans le cadre d'application de la question ***1.c)***.
  × Tout d'abord, comme $X_n(\Omega) = [\![0, n]\!]$ et $H_n(\Omega) = \{0, n\}$, alors : $(X_n - H_n)(\Omega) \subset [\![-n, n]\!]$.
  × Ensuite, la v.a.r. $\mathrm{e}^{-t\,H_n}$ admet une espérance en tant que v.a.r. finie (car H_n est une v.a.r. finie). Ainsi, par théorème de transfert, soit $t \in \mathbb{R}$:

$$\mathbb{E}\left(\mathrm{e}^{-t\,H_n}\right) = \mathrm{e}^{-0\,t} \, \mathbb{P}([H_n = 0]) + \mathrm{e}^{-nt} \, \mathbb{P}([H_n = n]) = \frac{1}{2} + \mathrm{e}^{-nt} \, \frac{1}{2} = \frac{1}{2} \left(1+\mathrm{e}^{-nt}\right)$$

  Ainsi :

$$\begin{aligned} M_{Y_n}(t) &= \frac{1}{2^n} \, (1+\mathrm{e}^t)^n \times \frac{1}{2} \, (1+\mathrm{e}^{-t})^n \\ &= M_{X_n}(t) \times \mathbb{E}(\mathrm{e}^{-t\,H_n}) \\ &= \mathbb{E}(\mathrm{e}^{t\,X_n}) \, \mathbb{E}(\mathrm{e}^{-t\,H_n}) \\ &= \mathbb{E}(\mathrm{e}^{t\,X_n} \, \mathrm{e}^{-t\,H_n}) && \textit{(car les v.a.r. } e^{t\,X_n} \textit{ et } e^{-t\,H_n} \textit{ sont indépendantes)} \\ &= M_{X_n-H_n}(t) \end{aligned}$$

  D'où : $M_{Y_n} = M_{X_n-H_n}$

  $$\boxed{\text{D'après } \textbf{\textit{1.c)}}, \text{ on en déduit que les v.a.r. } Y_n \text{ et } X_n - H_n \text{ suivent la même loi.}}$$

Commentaire

Les termes « en utilisant ... » de cette question font hésiter quant à la nécessité de démontrer l'égalité énoncée. Dans le doute, on le démontre.

□

Partie II. Propriétés générales des fonctions génératrices des cumulants et quelques exemples

5. Soit X une variable aléatoire et $\mathcal{D}_X$ le domaine de définition de la fonction K_X.

a) Donner la valeur de $K_X(0)$.

Démonstration.

- Tout d'abord, la v.a.r. $e^{0\,X} = 1$ (v.a.r. constante égale à 1) est une v.a.r. finie. Elle admet donc une espérance. Donc $M_X(0)$ est bien défini. De plus :

$$M_X(0) \;=\; \mathbb{E}\left(e^{0\,X}\right) \;=\; \mathbb{E}(1) \;=\; 1$$

- Ainsi, $K_X(0)$ est bien défini et :

$$K_X(0) \;=\; \ln\big(M_X(0)\big) \;=\; \ln(1) \;=\; 0$$

$$\boxed{K_X(0) = 0}$$

□

b) Soit $(a, b) \in \mathbb{R}^2$ et $Y = a\,X + b$. Justifier pour tout réel t pour lequel $a\,t$ appartient à $\mathcal{D}_X$, l'égalité :

$$K_Y(t) \;=\; b\,t + K_X(a\,t)$$

Démonstration.
Soit $t \in \mathbb{R}$ tel que $a\,t \in \mathcal{D}_X$.

- Comme $a\,t \in \mathcal{D}_X$, alors $K_X(a\,t)$ est bien défini, donc $M_X(a\,t)$ également et : $M_X(a\,t) > 0$.
- De plus :

$$M_X(a\,t) \;=\; \mathbb{E}\left(e^{a\,t}\right) \;=\; \mathbb{E}\left(e^{t\,(a\,X)}\right) \;=\; \mathbb{E}\left(e^{t\,(a\,X+b)-b\,t}\right) \;=\; \mathbb{E}\left(e^{t\,Y}\,e^{-b\,t}\right) \;=\; e^{-b\,t}\,\mathbb{E}\left(e^{t\,Y}\right)$$

Or, par définition de $M_Y(t)$ (si cette quantité existe) : $M_Y(t) = \mathbb{E}\left(e^{t\,T}\right)$.
D'après le calcul précédent, on en déduit que le réel $M_Y(t)$ est bien défini et :

$$M_Y(t) \;=\; e^{b\,t}\;M_X(a\,t)$$

- Comme : $M_X(a\,t) > 0$ et $e^{b\,t} > 0$, alors : $M_Y(t) = e^{b\,t}\;M_X(a\,t) > 0$.
On en déduit que $K_Y(t)$ est bien défini. Et enfin :

$$K_Y(t) \;=\; \ln\big(M_Y(t)\big) \;=\; \ln\left(e^{b\,t}\;M_X(a\,t)\right) \;=\; b\,t + \ln\big(M_X(a\,t)\big) \;=\; b\,t + K_X(a\,t)$$

$$\boxed{\text{Pour tout } t \in \mathbb{R} \text{ tel que } a\,t \in \mathcal{D}_X \,:\, K_Y(t) = b\,t + K_X(a\,t).}$$

Commentaire

On notera que la difficulté de cette question ne réside pas dans la démonstration de la relation entre $K_Y(t)$ et $K_X(a\,t)$ mais bien dans la démonstration de l'existence de tous les objets manipulés.

□

c) On suppose ici que les variables aléatoires X et $-X$ suivent la même loi.
Que peut-on dire dans ce cas des cumulants d'ordre impair de la variables aléatoire X ?

Démonstration.

- Soit $t \in \mathcal{D}_X$.
  Les v.a.r. X et $-X$ ont même loi. On en déduit que les v.a.r. $\mathrm{e}^{t\,X}$ et $\mathrm{e}^{-t\,X}$ ont même loi.
  Or, comme $t \in \mathcal{D}_X$, $K_X(t)$ existe, donc $\mathrm{e}^{t\,X}$ admet une espérance $(M_X(t))$ et : $M_X(t) > 0$.
  On en déduit que la v.a.r. $\mathrm{e}^{-t\,X}$ admet une espérance $(M_{-X}(t) = M_X(-t))$ et : $M_X(-t) > 0$.
  D'où : $-t \in \mathcal{D}_X$.

  $$\boxed{\text{On en déduit que, pour tout } t \in \mathcal{D}_X, \text{ on a : } -t \in \mathcal{D}_X.}$$

- Soit $t \in \mathcal{D}_X$, alors $-t \in \mathcal{D}_X$.
  - Ainsi, d'après la question précédente (appliquée à $a = -1$ et $b = 0$) :

    $$K_{-X}(t) \ = \ 0 \times t + K_X(-t) \ = \ K_X(-t)$$

  - De plus, comme X et $-X$ ont même loi (on rappelle que $M_X(t)$ et $M_{-X}(t)$ sont bien définis car $t \in \mathcal{D}_X$ et $-t \in \mathcal{D}_X$) :

    $$M_{-X}(t) \ = \ \mathbb{E}\left(\mathrm{e}^{t\,(-X)}\right) \ = \ \mathbb{E}\left(\mathrm{e}^{t\,X}\right) \ = \ M_X(t)$$

    Ainsi : $K_{-X}(t) \ = \ \ln\big(M_{-X}(t)\big) \ = \ \ln\big(M_X(t)\big) \ = \ K_X(t)$.

  On en déduit : $K_X(t) = K_{-X}(t) = K_X(-t)$.

  $$\boxed{\text{Ainsi, si les v.a.r. } X \text{ et } -X \text{ suivent la même loi : } \forall t \in \mathcal{D}_X,\ K_X(t) = K_X(-t).}$$

- Supposons maintenant que la fonction K_X est de classe $\mathcal{C}^\infty$ sur $\mathcal{D}_X$. Soit $p \in \mathbb{N}$:
  × tout d'abord : $Q^{(2p+1)}(X) = K_X^{(2p+1)}(0)$,
  × ensuite : $\forall t \in \mathcal{D}_X$, $K_X(t) = K_X(-t)$.
  Alors, par récurrence immédiate : $\forall n \in \mathbb{N}$, $\forall t \in \mathcal{D}_X$, $K_X^{(n)}(t) = (-1)^n\, K_X^{(n)}(-t)$.
  En particulier, pour tout $t \in \mathcal{D}_X$:

  $$K_X^{(2p+1)}(t) \ = \ (-1)^{2p+1}\, K_X^{(2p+1)}(-t) \ = \ -K_X^{(2p+1)}(-t)$$

  Comme $0 \in \mathcal{D}_X$ (d'après ***5.a)***), on en déduit :

  $$K_X^{(2p+1)}(0) = -K^{(2p+1)}(0) \ \Leftrightarrow \ 2\,K_X^{(2p+1)}(0) = 0 \ \Leftrightarrow \ K_X^{(2p+1)}(0) = 0$$

$$\boxed{\text{Ainsi, si } X \text{ et } -X \text{ suivent la même loi, sous réserve d'existence : } \forall p \in \mathbb{N},\ Q^{(2p+1)}(X) = 0.}$$

Commentaire

- Remarquons que si K_X n'est pas dérivable sur $\mathcal{D}_X$ alors on ne peut rien dire des cumulants de X, puisque ces derniers n'existent pas.
- Dans le programme ECE, on trouve la propriété :

Les v.a.r. X et Y ont même loi	$\Leftrightarrow$	Les v.a.r. X et Y ont même fonction de répartition

- On peut alors démontrer que si X et Y sont des v.a.r. discrètes (resp. à densité), on a :

$$\left.\begin{array}{l}\bullet \text{ Les v.a.r. } X \text{ et } Y \text{ ont même loi} \\ \bullet \text{ Les v.a.r. } X \text{ et } Y \text{ admettent un} \\ \text{ moment d'ordre } n \in \mathbb{N}^* \end{array}\right\} \Rightarrow \mathbb{E}\big(X^n\big) = \mathbb{E}\big(Y^n\big)$$

  Cette dernière propriété est aussi vérifiée pour les v.a.r. quelconques mais n'est pas explicitement écrite dans ce cas précis. Toutefois, on peut considérer qu'on y a accès puisque le programme précise : « on admettra que les propriétés opératoires usuelles de l'espérance et de la variance se généralisent aux variables aléatoires quelconques ».
- Enfin, on utilise dans cette question la propriété :

$$X \text{ et } Y \text{ ont même loi} \quad \Rightarrow \quad f(X) \text{ et } f(Y) \text{ ont même loi}$$

  où f est une fonction (ici $x \mapsto \mathrm{e}^{t\,x}$).
  Cette propriété est facile à démontrer dans le cas de v.a.r. discrètes ou à densité. On l'admettra pour le cas de v.a.r. quelconques.

$\square$

6. Soit X et Y deux variables aléatoires indépendantes et $\mathcal{D}_X$ et $\mathcal{D}_Y$ les domaines de définition respectifs des fonctions K_X et K_Y.

a) Monter que pour tout réel t appartenant à la fois à $\mathcal{D}_X$ et $\mathcal{D}_Y$, on a : $K_{X+Y}(t) = K_X(t) + K_Y(t)$.

Démonstration.
Soit $t \in \mathcal{D}_X \cap \mathcal{D}_Y$.

- Comme $t \in \mathcal{D}_X \cap \mathcal{D}_Y$, les quantités $K_X(t)$ et $K_Y(t)$ sont bien définies.
  En particulier, $M_X(t)$ et $M_Y(t)$ sont bien définies et : $M_X(t) > 0$ et $M_Y(t) > 0$.
  Ainsi, les v.a.r. $\mathrm{e}^{t\,X}$ et $\mathrm{e}^{t\,Y}$:
  - × admettent une espérance,
  - × sont indépendantes par lemme des coalitions, car X et Y sont indépendantes.

  On en déduit que la v.a.r. $\mathrm{e}^{t\,X} \times \mathrm{e}^{t\,Y} = \mathrm{e}^{t\,(X+Y)}$ admet une espérance.

  $\boxed{\text{Ainsi, la quantité } M_{X+Y}(t) \text{ est bien définie.}}$

Commentaire

- On utilise ici le fait que si deux v.a.r. U et V sont indépendantes et admettent une espérance alors, la v.a.r. produit UV admet une espérance donnée par : $\mathbb{E}\big(UV\big) = \mathbb{E}(U)\,\mathbb{E}(V)$.
  L'hypothèse d'indépendance est ici cruciale pour démontrer l'existence de l'espérance du produit et pour obtenir sa valeur.

Commentaire

- Dans le cas général, la v.a.r. produit UV admet une espérance si les v.a.r. U et V admettent **un moment d'ordre** 2.
- On peut se demander d'où provient cette hypothèse liée aux moments d'ordre 2.
  Elle est issue d'un théorème de domination. Détaillons ce point.
  Remarquons tout d'abord : $(U - V)^2 \geqslant 0$.
  On en déduit : $U^2 - 2\,UV + V^2 \geqslant 0$.
  Et, en réordonnant : $UV \leqslant \dfrac{1}{2}\,U^2 + \dfrac{1}{2}\,V^2$. Ou encore :
  $$0 \ \leqslant\ |UV| \ \leqslant\ \frac{1}{2}\,U^2 + \frac{1}{2}\,V^2$$
  Comme U et V admettent un moment d'ordre 2, la v.a.r. $\frac{1}{2}\,U^2 + \frac{1}{2}\,V^2$ admet une espérance comme combinaison linéaire de v.a.r. qui en admettent une.
  Ainsi, par théorème de domination (présenté seulement dans le programme ECS), la v.a.r. $|UV|$ admet une espérance. Il en est de même de la v.a.r. UV.

- De plus :

$$\begin{array}{rcll} M_{X+Y}(t) & = & \mathbb{E}\left(\mathrm{e}^{t\,(X+Y)}\right) & \\ & = & \mathbb{E}\left(\mathrm{e}^{t\,X}\ \mathrm{e}^{t\,Y}\right) & \\ & = & \mathbb{E}\left(\mathrm{e}^{t\,X}\right)\ \mathbb{E}\left(\mathrm{e}^{t\,Y}\right) & \textit{(car les v.a.r. } \mathrm{e}^{t\,X} \textit{ et } \mathrm{e}^{t\,Y} \textit{ sont indépendantes)} \\ & = & M_X(t)\ M_Y(t) & \end{array}$$

- Enfin, comme $M_X(t) > 0$ et $M_Y(t) > 0$, alors : $M_{X+Y}(t) > 0$. Donc $K_{X+Y}(t)$ est bien définie.
  On obtient :

$$K_{X+Y}(t) \ = \ \ln\big(M_{X+Y}(t)\big) \ = \ \ln\big(M_X(t)\,M_Y(t)\big) \ = \ \ln\big(M_X(t)\big) + \ln\big(M_Y(t)\big) \ = \ K_X(t) + K_Y(t)$$

$$\boxed{\forall t \in \mathcal{D}_X \cap \mathcal{D}_Y,\ K_{X+Y}(t) = K_X(t) + K_Y(t)}$$

□

b) En déduire une relation entre les cumulants des variables aléatoires X, Y et $X+Y$.

Démonstration.
Soit $p \in \mathbb{N}^*$.
Si K_X est de classe $\mathcal{C}^p$ sur $\mathcal{D}_X$ et K_Y est de classe $\mathcal{C}^p$ sur $\mathcal{D}_Y$, alors, d'après la question précédente, K_{X+Y} est de classe $\mathcal{C}^p$ sur $\mathcal{D}_X \cap \mathcal{D}_Y$ et :

$$\forall t \in \mathcal{D}_X \cap \mathcal{D}_Y, \quad K_{X+Y}^{(p)}(t) \ = \ K_X^{(p)}(t) + K_X^{(p)}(t)$$

Or $0 \in \mathcal{D}_X \cap \mathcal{D}_Y$. Donc :

$$Q_p(X+Y) \ = \ K_{X+Y}^{(p)}(0) \ = \ K_X^{(p)}(0) + K_Y^{(p)}(0) \ = \ Q_p(X) + Q_p(Y)$$

$$\boxed{\begin{array}{c}\text{Finalement, pour tout } p \in \mathbb{N}^*, \text{ si } Q_p(X) \text{ et } Q_p(Y) \text{ sont bien définis, alors :} \\ Q_p(X+Y) \ = \ Q_p(X) + Q_p(Y).\end{array}}$$

□

7. Soit U une variable aléatoire suivant la loi uniforme sur l'intervalle $[0,1]$.

a) Montrer que la fonction M_U est définie sur $\mathbb{R}$ et donnée par : $\forall t \in \mathbb{R}$, $M_U(t) = \begin{cases} \dfrac{e^t - 1}{t} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0 \end{cases}$.

Démonstration.

- Soit $t \in \mathbb{R}$.
  Le réel $M_U(t)$ existe si et seulement si la v.a.r. $e^{t\,U}$ admet une espérance.
  Par théorème de transfert, la v.a.r. $e^{t\,U}$ admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} e^{t\,x} f_U(x)\ dx$ est absolument convergente.
  Les fonctions $x \mapsto e^{t\,x}$ et f_U étant à valeurs positives sur $\mathbb{R}$ (f_U est une densité de probabilité), cela revient à démonter que cette intégrale est convergente.
- De plus, la fonction f_u est nulle en dehors de $[0,1]$, donc :

$$\int_{-\infty}^{+\infty} e^{t\,x} f_U(x)\ dx \ = \ \int_0^1 e^{t\,x} f_U(x)\ dx$$

- La fonction $x \mapsto e^{t\,x} f_U(x)$ est continue par morceaux sur le segment $[0,1]$. On en déduit que l'intégrale $\displaystyle\int_0^1 e^{t\,x} f_U(x)\ dx$ est bien définie.
  Ainsi, pour tout $t \in \mathbb{R}$, la v.a.r. $e^{t\,U}$ admet une espérance.

$$\boxed{\text{La fonction } M_U \text{ est donc définie sur } \mathbb{R}.}$$

- Soit $t \in \mathbb{R}$. Deux cas se présentent :
  × si $t = 0$, alors d'après la question ***5.a)*** : $M_U(0) = 1$.
  × si $t \neq 0$, alors :

$$\begin{array}{rcll} M_U(t) & = & \mathbb{E}\left(e^{t\,U}\right) & \\ & = & \displaystyle\int_0^1 e^{t\,x} f_U(x)\ dx & \textit{(après les points précédents)} \\ & = & \displaystyle\int_0^1 e^{t\,x}\ dx & \\ & = & \left[\dfrac{1}{t}\ e^{t\,x}\right]_0^1 & \textit{(car } t \neq 0\textit{)} \\ & = & \dfrac{1}{t}\left(e^t - 1\right) & \end{array}$$

$$\boxed{\text{Finalement : } M_U : t \mapsto \begin{cases} \dfrac{e^t - 1}{t} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0 \end{cases}.}$$

□

b) Calculer la dérivée de la fonction M_U en tout point $t \neq 0$.

Démonstration.

- La fonction M_U est dérivable sur $]-\infty, 0[$ et sur $]0, +\infty[$ en tant que quotient $\dfrac{f_1}{f_2}$ avec :
  - × $f_1 : t \mapsto \mathrm{e}^t - 1$ dérivable sur $]-\infty, 0[$ et sur $]0, +\infty[$,
  - × $f_2 : t \mapsto t$ dérivable sur $]-\infty, 0[$ et sur $]0, +\infty[$ et qui ne s'annule pas sur cet intervalle.
- Soit $t \in\,]-\infty, 0[\cup]0, +\infty[$.

$$M_U'(t) \;=\; \frac{\mathrm{e}^t \times t - (\mathrm{e}^t - 1) \times 1}{t^2}$$

$$\boxed{\forall t \in\,]-\infty, 0[\cup]0, +\infty[,\ M_U(t) = \frac{t\,\mathrm{e}^t - \mathrm{e}^t + 1}{t^2}}$$

□

c) Trouver la limite du quotient $\dfrac{M_U(t) - 1}{t}$ lorsque t tend vers 0.

Démonstration.

- Soit $t \in \mathbb{R}^*$.

$$\frac{M_U(t) - 1}{t} = \frac{\frac{\mathrm{e}^t - 1}{t} - 1}{t} = \frac{\frac{\mathrm{e}^t - 1 - t}{t}}{t} = \frac{\mathrm{e}^t - 1 - t}{t^2}$$

- Déterminons, si elle existe, $\lim\limits_{t \to 0} \dfrac{\mathrm{e}^t - 1 - t}{t^2}$.
  - × Tout d'abord : $\mathrm{e}^t = 1 + t + \dfrac{t^2}{2} + \underset{t \to 0}{o}(t^2)$. Ainsi : $\mathrm{e}^t - 1 - t = \dfrac{t^2}{2} + \underset{t \to 0}{o}(t^2)$. D'où :

$$\mathrm{e}^t - 1 - t \underset{t \to 0}{\sim} \frac{t^2}{2}$$

  - × On en déduit :

$$\frac{\mathrm{e}^t - 1 - t}{t^2} \underset{t \to 0}{\sim} \frac{\frac{t^2}{2}}{t^2} = \frac{1}{\cancel{t^2}} \frac{\cancel{t^2}}{2} = \frac{1}{2}$$

$$\boxed{\text{On en déduit : } \lim_{t \to 0} \frac{M_U(t) - 1}{t} = \frac{1}{2}.}$$

□

d) Montrer que la fonction M_U est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

Démonstration.

- D'après la question précédente : $\lim\limits_{t \to 0} \dfrac{M_U(t) - M_U(0)}{t - 0} \;=\; \dfrac{1}{2}$.

$$\boxed{\text{On en déduit que la fonction } M_U \text{ est dérivable en 0 et } M_U'(0) = \frac{1}{2}.}$$

- De plus, d'après la question ***7.b)***, la fonction M_U est dérivable sur $]-\infty, 0[$ et sur $]0, +\infty[$.

$$\boxed{\text{On en déduit que la fonction } M_U \text{ est dérivable sur } \mathbb{R}.}$$

- La fonction M_U est de classe $\mathcal{C}^1$ sur $]-\infty, 0[$ et sur $]0, +\infty[$ avec des arguments similaires à ceux de la dérivabilité sur ces intervalles (question ***7.b)***).
- On cherche enfin à montrer que la fonction M_U est de classe $\mathcal{C}^1$ en 0, c'est-à-dire que la fonction M_U' est continue en 0. On rappelle :

$$M_U' : t \mapsto \begin{cases} \dfrac{t\,\mathrm{e}^t - \mathrm{e}^t + 1}{t^2} & \text{si } t \neq 0 \\[2ex] \dfrac{1}{2} & \text{si } t = 0 \end{cases}$$

On cherche donc à montrer : $\lim\limits_{t \to 0} \dfrac{t\,\mathrm{e}^t - \mathrm{e}^t + 1}{t^2} = \dfrac{1}{2}$.

× On sait déjà : $\mathrm{e}^t = 1 + t + \dfrac{t^2}{2} + \underset{t \to 0}{o}(t^2)$. Ainsi :

$$t\,\mathrm{e}^t \;=\; t\left(1 + t + \frac{t^2}{2} + \underset{t \to 0}{o}(t^2)\right) \;=\; t + t^2 + \frac{t^3}{2} + \underset{t \to 0}{o}(t^3)$$

On en déduit :

$$\begin{aligned} t\,\mathrm{e}^t - \mathrm{e}^t + 1 &= \cancel{t} + t^2 + \frac{t^3}{2} + \underset{t \to 0}{o}(t^3) - \left(\cancel{1} + \cancel{t} + \frac{t^2}{2} + \underset{t \to 0}{o}(t^2)\right) + \cancel{1} \\ &= \frac{t^2}{2} + \underset{t \to 0}{o}(t^2) \qquad \textit{(car } t^3 = \underset{t \to 0}{o}(t^2)\textit{)} \end{aligned}$$

Ainsi : $t\,\mathrm{e}^t - \mathrm{e}^t + 1 \underset{t \to 0}{\sim} \dfrac{t^2}{2}$.

× On obtient :

$$M_U'(t) \;=\; \frac{t\,\mathrm{e}^t - \mathrm{e}^t + 1}{t^2} \underset{t \to 0}{\sim} \frac{\frac{\cancel{t^2}}{2}}{\cancel{t^2}} \;=\; \frac{1}{2}$$

On en déduit : $\lim\limits_{t \to 0} M_U'(t) \;=\; \dfrac{1}{2} \;=\; M_U'(0)$.
La fonction M_U' est donc continue en 0.

$$\boxed{\text{Finalement, la fonction } M_U \text{ est de classe } \mathcal{C}^1 \text{ sur } \mathbb{R}.}$$

Commentaire

- On pouvait également déterminer le DL à l'ordre 2 « $t\,\mathrm{e}^t - \mathrm{e}^t + 1 = \dfrac{t^2}{2} + \underset{t \to 0}{o}(t^2)$ » en appliquant la formule de Taylor-Young à la fonction $t \mapsto t\,\mathrm{e}^t - \mathrm{e}^t + 1$ (qui est bien de classe $\mathcal{C}^2$ sur $\mathbb{R}$).
- L'utilisation de la formule de Taylor-Young serait un peu plus cohérente au regard du programme d'ECE. On ne privilégie cependant pas cette méthode ici puisqu'elle est bien plus chronophage que celle présentée.

□

8. Soit α et β deux réels tels que $\alpha < \beta$.
Dans cette question, on note X une variable aléatoire qui suit la loi uniforme sur l'intervalle $[\alpha, \beta]$.

a) Exprimer K_X en fonction de M_U, où la variable aléatoire U a été définie dans la question ***7.***

Démonstration.

- Montrons que la fonction K_U est bien définie, *i.e.* : $\forall t \in \mathbb{R}$, $M_U(t) > 0$.
Soit $t \in \mathbb{R}$. Trois cas se présentent :
   × si $t \in]-\infty, 0[$:
$$\begin{aligned} M_u(t) > 0 \quad &\Leftrightarrow \quad \frac{e^t - 1}{t} > 0 \\ &\Leftrightarrow \quad e^t - 1 < 0 \qquad \textit{(car } t < 0\textit{)} \\ &\Leftrightarrow \quad e^t < 1 \\ &\Leftrightarrow \quad t < 0 \qquad \textit{(par stricte croissance de la fonction } \ln \textit{ sur }]0, +\infty[\textit{)} \end{aligned}$$
   La dernière inégalité est vraie, donc, par équivalence, la première aussi.
   × si $t = 0$, alors : $M_U(0) = 1 > 0$.
   × si $t \in]0, +\infty[$, alors, avec un raisonnement similaire au cas $t \in]-\infty, 0[$, on obtient également : $M_U(t) > 0$.

$$\boxed{\text{La fonction } K_U \text{ est bien définie sur } \mathbb{R}.}$$

Commentaire

- L'esprit du sujet est ici de faire l'étude des fonctions M_X et K_X dans le cas particulier de certaines lois usuelles. C'est pourquoi on exploite l'expression explicite de M_U pour démontrer l'existence de K_U.
- On pourrait en fait démontrer dans un cadre très général que si $M_X(t)$ est bien définie, alors, comme la v.a.r. e^{tX} est à valeurs **strictement** positives : $M_X(t) > 0$.
On démontrera cette implication dans la partie III qui, elle traite du cas général.

- De plus, rappelons :
$$U \hookrightarrow \mathcal{U}([0,1]) \quad \Leftrightarrow \quad Y = (\beta - \alpha)\, U + \alpha \hookrightarrow \mathcal{U}([\alpha, \beta])$$
Ainsi, en notant $Y = (\beta - \alpha)\, U + \alpha$, on obtient que les v.a.r. Y et X ont même loi.
- Soit $t \in \mathbb{R}$, alors $(\beta - \alpha)\, t \in \mathbb{R}$. On en déduit, d'après la question ***5.b)***, que $K_Y(t)$ existe et :
$$K_Y(t) \;=\; \alpha\, t + K_U\big((\beta - \alpha)\, t\big) \;=\; \alpha\, t + \ln\Big(M_U\big((\beta - \alpha)\, t\big)\Big)$$
(on rappelle que, comme la fonction K_Y est bien définie sur $\mathbb{R}$, la fonction M_Y aussi et : $\forall t \in \mathbb{R}$, $M_Y(t) > 0$)
- Enfin, comme les v.a.r. X et Y ont même loi, alors les v.a.r. e^{tX} et e^{tY} ont même loi.
Or, comme la fonction M_Y est bien définie sur $\mathbb{R}$, la v.a.r. e^{tY} admet une espérance strictement positive. Ainsi, la v.a.r. e^{tX} admet la même espérance strictement positive, *i.e.* la fonction M_X est définie sur $\mathbb{R}$ et : $\forall t \in \mathbb{R}$, $M_X(t) = M_Y(t) > 0$.

$$\boxed{\begin{array}{c}\text{On en déduit que la fonction } K_X \text{ est définie sur } \mathbb{R} \text{ et :} \\ \forall t \in \mathbb{R}, \quad K_X(t) \;=\; \alpha\, t + \ln\Big(M_U\big((\beta - \alpha)\, t\big)\Big). \end{array}}$$

□

b) Justifier que la fonction K_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et établir l'égalité : $Q_1(X) = \mathbb{E}(X)$.

Démonstration.

- Tout d'abord, la fonction $\widetilde{M_U} : t \mapsto M_u\big((\beta - \alpha)\, t\big)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ en tant que composée de fonctions de classe $\mathcal{C}^1$ sur $\mathbb{R}$ (on rappelle que M_U est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ d'après ***7.d)***).
  De plus, comme $M_U(]-\infty, +\infty[) \subset]0, +\infty[$ (démontré dans la question précédente), pour tout $t \in \mathbb{R}$: :
  $$\widetilde{M_U}(t) = M_U\big((\beta - \alpha)\, t\big) > 0$$
- Ensuite, la fonction $t \mapsto \ln\Big(M_u\big((\beta - \alpha)\, t\big)\Big)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ car la composée $\ln \circ \tilde{M}_U$ de :
  - × $\tilde{M}_U : t \mapsto M_U\big((\beta - \alpha)\, t\big)$ qui :
    - est de classe $\mathcal{C}^1$ sur $\mathbb{R}$,
    - vérifie : $\tilde{M}_U(]-\infty, +\infty[) \subset]0, +\infty[$
  - × $\ln$ qui est de classe $\mathcal{C}^1$ sur $]0, +\infty[$.

On en déduit que la fonction K_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ en tant que somme de fonctions de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

- Par définition de $Q_1(X)$, on a : $Q_1(X) = K'_X(0)$.
  Or, d'après la question précédente, pour tout $t \in \mathbb{R}$:
  $$K'_X(t) \ = \ \alpha + \frac{(\beta - \alpha)\, M'_U\big((\beta - \alpha)\, t\big)}{M_U\big((\beta - \alpha)\, t\big)}$$
  On en déduit :
  $$\begin{aligned} Q_1(X) \ &= \ K'_X(0) \ = \ \alpha + \frac{(\beta - \alpha)\, M'_U(0)}{M_U(0)} \\ &= \ \alpha + \frac{(\beta - \alpha)\ \frac{1}{2}}{1} \qquad \textit{(d'après } \textbf{\textit{7.a)}} \textit{ et } \textbf{\textit{7.d)}}\textit{)} \\ &= \ \alpha + \frac{\beta - \alpha}{2} \\ &= \ \frac{\alpha + \beta}{2} \end{aligned}$$

Finalement, comme $X \hookrightarrow \mathcal{U}([\alpha, \beta])$: $\mathbb{E}(X) = \dfrac{\alpha + \beta}{2} = Q_1(X)$. □

9. Soit un réel $\lambda > 0$ et soit T une variable aléatoire qui suit la loi de Poisson de paramètre λ.

a) Déterminer les fonctions M_T et K_T.

Démonstration.

- Soit $t \in \mathbb{R}$. Tout d'abord, la quantité $M_T(t)$ est bien définie si la v.a.r. $\mathrm{e}^{t\,T}$ admet une espérance.
  Or, par théorème de transfert, la v.a.r. $\mathrm{e}^{t\,T}$ admet une espérance si et seulement si la série $\sum\limits_{n \geqslant 0} \mathrm{e}^{t\,n}\ \mathbb{P}([T = n])$ est absolument convergente.
  Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.

- Soit $N \in \mathbb{N}$.

$$\begin{array}{rcll} \sum\limits_{k=0}^{N} \mathrm{e}^{k\,t}\ \mathbb{P}([T=k]) & = & \sum\limits_{k=0}^{N} \mathrm{e}^{k\,t}\ \dfrac{\lambda^k}{k!}\ \mathrm{e}^{-\lambda} & \textit{(car } T \hookrightarrow \mathcal{P}(\lambda)\textit{)} \\ & = & \mathrm{e}^{-\lambda}\ \sum\limits_{k=0}^{N} \dfrac{\left(\mathrm{e}^{t}\right)^k\ \lambda^k}{k!} & \\ & = & \mathrm{e}^{-\lambda}\ \sum\limits_{k=0}^{N} \dfrac{\left(\lambda\,\mathrm{e}^{t}\right)^k}{k!} & \end{array}$$

On reconnaît la somme partielle d'ordre N de la série exponentielle $\sum\limits_{n\geqslant 0} \dfrac{x^n}{n!}$ en $x = \lambda\,\mathrm{e}^{t}$. Cette série est convergente et :

$$\sum_{k=0}^{N} \frac{\left(\lambda\,\mathrm{e}^{t}\right)^k}{k!} \underset{N\to+\infty}{\longrightarrow} \sum_{k=0}^{+\infty} \frac{\left(\lambda\,\mathrm{e}^{t}\right)^k}{k!} = \exp\left(\lambda\,\mathrm{e}^{t}\right)$$

On en déduit que la série $\sum\limits_{n\geqslant 0} \mathrm{e}^{t\,n}\ \mathbb{P}([T=n])$ converge. Ainsi, pour tout $t \in \mathbb{R}$, la v.a.r. $\mathrm{e}^{t\,T}$ admet une espérance.

$$\boxed{\text{On en déduit que la fonction } M_T \text{ est définie sur } \mathbb{R}.}$$

- De plus, soit $t \in \mathbb{R}$:

$$M_T(t) \ = \ \mathbb{E}\left(\mathrm{e}^{t\,T}\right) \ = \ \mathrm{e}^{-\lambda}\ \exp\left(\lambda\,\mathrm{e}^{t}\right) \ = \ \exp\left(-\lambda + \lambda\,\mathrm{e}^{t}\right) \ = \ \exp\Big(\lambda\left(\mathrm{e}^{t}-1\right)\Big)$$

$$\boxed{M_T : t \mapsto \exp\Big(\lambda\left(\mathrm{e}^{t}-1\right)\Big)}$$

- On a bien, pour tout $t \in \mathbb{R}$: $M_T(t) = \exp\Big(\lambda\left(\mathrm{e}^{t}-1\right)\Big) > 0$.

$$\boxed{\text{La fonction } K_T \text{ est donc définie sur } \mathbb{R}.}$$

- Soit $t \in \mathbb{R}$.

$$K_T(t) \ = \ \ln\big(M_T(t)\big) \ = \ \ln\Big(\exp\Big(\lambda\left(\mathrm{e}^{t}-1\right)\Big)\Big) \ = \ \lambda\left(\mathrm{e}^{t}-1\right)$$

$$\boxed{K_T : t \mapsto \lambda\left(\mathrm{e}^{t}-1\right)}$$

□

b) En déduire les cumulants de T.

Démonstration.

- Tout d'abord, la fonction K_T est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ en tant que transformée affine de la fonction exp de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

$$\boxed{\text{Ainsi, la v.a.r. } T \text{ admet des cumulants à tout ordre.}}$$

- Ensuite, pour tout $t \in \mathbb{R}$: $K_T'(t) = \lambda\,\mathrm{e}^{t}$.
On en déduit, par récurrence immédiate, que pour tout $n \in \mathbb{N}^*$: $\forall t \in \mathbb{R}$, $K_T^{(n)}(t) = \lambda\,\mathrm{e}^{t}$.

$$\boxed{\text{Ainsi, pour tout } n \in \mathbb{N}^* : Q_n(X) = K_T^{(n)}(0) = \lambda.}$$

□

10. Soit Z une variable aléatoire qui suit la loi normale centrée réduite.

a) Justifier pour tout $t \in \mathbb{R}$, la convergence de l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \exp\left(t\,x - \frac{x^2}{2}\right)\, dx$.

Démonstration.
Soit $t \in \mathbb{R}$.

- L'intégrale impropre $\displaystyle\int_{-\infty}^{+\infty} \exp\left(t\,x - \frac{x^2}{2}\right)\, dx$ est convergente si et seulement si :
  - l'intégrale impropre $\displaystyle\int_{-\infty}^{0} \exp\left(t\,x - \frac{x^2}{2}\right)\, dx$ est convergente.
  - et l'intégrale impropre $\displaystyle\int_{0}^{+\infty} \exp\left(t\,x - \frac{x^2}{2}\right)\, dx$ est convergente.
- Démontrons tout d'abord la convergence de l'intégrale impropre $\displaystyle\int_{0}^{+\infty} \exp\left(t\,x - \frac{x^2}{2}\right)\, dx$.
  La fonction $h_t : x \mapsto \exp\left(t\,x - \dfrac{x^2}{2}\right)$ est continue sur $[0, +\infty[$.
  Ainsi, l'intégrale $\displaystyle\int_{0}^{+\infty} h_t(x)\, dx$ est impropre seulement en $+\infty$.
  - Tout d'abord, la fonction h_t est continue sur le **segment** $[0, 1]$.
    On en déduit que l'intégrale $\displaystyle\int_{0}^{1} h_t(x)\, dx$ est bien définie.
  - Par ailleurs :
    × $\forall x \in [1, +\infty[$, $\exp\left(t\,x - \dfrac{x^2}{2}\right) \geqslant 0$ et $\dfrac{1}{x^2} \geqslant 0$.
    × $\exp\left(t\,x - \dfrac{x^2}{2}\right) \underset{x \to +\infty}{=} o\left(\dfrac{1}{x^2}\right)$
    × $\displaystyle\int_{1}^{+\infty} \frac{1}{x^2}\, dx$ est une intégrale de Riemann, impropre en $+\infty$, d'exposant 2 (> 1).
    Elle est donc convergente.

    Par critère de négligeabilité d'intégrales généralisées de fonctions continues positives, l'intégrale $\displaystyle\int_{1}^{+\infty} h_t(x)\, dx$ est convergente.

$$\boxed{\text{Ainsi, pour tout } t \in \mathbb{R}, \text{ l'intégrale impropre } \displaystyle\int_{0}^{+\infty} h_t(x)\, dx \text{ est convergente.}}$$

- Il reste à démontrer la convergence de l'intégrale impropre $\displaystyle\int_{0}^{+\infty} \exp\left(t\,x - \frac{x^2}{2}\right)\, dx$.
  La fonction h_t est continue sur $]-\infty, 0]$ donc cette intégrale est impropre seulement en $-\infty$.
  En effectuant le changement de variable affine $\boxed{u = -x}$, on obtient :

$$\begin{aligned}
\int_{-\infty}^{0} \exp\left(t\,x - \frac{x^2}{2}\right)\, dx &= \int_{0}^{+\infty} \exp\left(-t\,u - \frac{(-u)^2}{2}\right)(-\, du) \\
&= -\int_{0}^{+\infty} \exp\left(-t\,u - \frac{u^2}{2}\right)\, du \;=\; -\int_{0}^{+\infty} f_{-t}(u)\, du
\end{aligned}$$

En appliquant le résultat précédent en $-t \in \mathbb{R}$, on conclut que l'intégrale impropre $\displaystyle\int_0^{+\infty} f_{-t}(u)\, du$ est convergente.

$$\boxed{\text{Ainsi, pour tout } t \in \mathbb{R}, \text{ l'intégrale impropre } \int_{-\infty}^{0} h_t(x)\, dx \text{ est convergente.}}$$

$$\boxed{\text{Pour tout } t \in \mathbb{R}, \text{ l'intégrale impropre } \int_{-\infty}^{+\infty} h_t(x)\, dx \text{ est bien convergente.}}$$

Commentaire

- Le programme officiel précise que « les changements de variables **non affines** ne seront pratiqués qu'avec des intégrales sur un segment ». Il est donc autorisé, sous réserve de convergence, d'effectuer un changement de variable affine sur une intégrale généralisée (ce qui est fait dans cette question).
- Ici, on pose le changement de variable affine $\boxed{u = -x}$. Il faut s'habituer à effectuer ce changement de variable classique à la volée. Rappelons comment le présenter formellement.

$$\left|\begin{array}{l} u = -x \text{ (et donc } x = -u) \\ \hookrightarrow du = -\, dx \quad \text{et} \quad dx = -\, du \\ \bullet\ x = -\infty \ \Rightarrow\ u = +\infty \\ \bullet\ x = 0 \ \Rightarrow\ u = 0 \end{array}\right.$$

  Ce changement de variable est valide car $\psi : u \mapsto -u$ est de classe $\mathcal{C}^1$ sur $[0, +\infty[$.

- La fonction intégrande $h_t : x \mapsto e^{t\,x - \frac{x^2}{2}}$:
  - × n'admet pas d'équivalent plus simple en $+\infty$,
  - × tend très rapidement vers 0 en $+\infty$.

  Du fait de ces deux points, on opte dans la démonstration pour un critère de négligeabilité. De manière informelle, la présence du terme $e^{-\frac{x^2}{2}}$ produit une convergence extrêmement rapide de h_t vers 0 en $+\infty$. L'intégrande h_t apparaît donc suffisamment petite en $+\infty$ pour que l'intégrale sur $[1, +\infty[$ associée soit convergente. Formellement, cette idée est concrétisée en comparant h_t à l'intégrande $x \mapsto \frac{1}{x^2}$, positive et dont l'intégrale sur $[1, +\infty[$ associée est convergente.

- La démonstration de la propriété $\exp\left(t\,x - \dfrac{x^2}{2}\right) \underset{x \to +\infty}{=} o\left(\dfrac{1}{x^2}\right)$ n'est pas forcément un attendu de la question. Prendre l'initiative de la comparaion à la fonction $x \mapsto \frac{1}{x^2}$ démontre la bonne compréhension des mécanismes en jeu. Il faut évidemment savoir comment faire cette démonstration dont on donne ci-dessous les détails :

$$\frac{e^{t\,x - \frac{x^2}{2}}}{\frac{1}{x^2}} = x^2\, e^{t\,x - \frac{x^2}{2}} = \frac{x^2}{e^{\frac{x^2}{4}}}\, e^{t\,x - \frac{x^2}{4}} \underset{x \to +\infty}{\longrightarrow} 0$$

  En effet :
  - × en posant $u = \frac{x^2}{4}$, on obtient : $\displaystyle\lim_{x \to +\infty} \frac{x^2}{e^{\frac{x^2}{4}}} = \lim_{u \to +\infty} 4\, \frac{u}{e^u} = 0$ *(par croissances comparées)*.
  - × $\displaystyle\lim_{x \to +\infty} e^{t\,x - \frac{x^2}{2}} = 0$.

□

b) Montrer que la fonction M_Z est définie sur $\mathbb{R}$ et donnée par : $\forall t \in \mathbb{R},\ M_Z(t) = \exp\left(\dfrac{t^2}{2}\right)$.

Démonstration.
Soit $t \in \mathbb{R}$.

- Le réel $M_Z(t)$ existe si et seulement si la v.a.r. $\mathrm{e}^{t\,Z}$ admet une espérance.
  Or, par théorème de transfert, la v.a.r. $\mathrm{e}^{t\,Z}$ admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{t\,x}\ f_Z(x)\ dx$ est absolument convergente.
  Les fonctions $x \mapsto \mathrm{e}^{t\,x}$ et f_Z étant à valeurs positives sur $\mathbb{R}$ (f_Z est une densité de probabilité), cela revient à démonter que cette intégrale est convergente.

  $$\boxed{\text{D'après la question précédente, cette intégrale est convergente pour tout } t \in \mathbb{R}. \quad \text{Ainsi, } M_Z \text{ est définie sur } \mathbb{R}.}$$

- Remarquons que pour tout $x \in \,]-\infty, +\infty[$:
  $$\mathrm{e}^{t\,x}\ f_Z(x) \ = \ \frac{1}{\sqrt{2\,\pi}}\,\mathrm{e}^{t\,x}\ \mathrm{e}^{-\frac{1}{2}\,x^2} \ = \ \frac{1}{\sqrt{2\,\pi}}\,\mathrm{e}^{-\frac{1}{2}\,(x^2 - 2\,t x)} \ = \ \frac{1}{\sqrt{2\,\pi}}\,\mathrm{e}^{-\frac{1}{2}\,(x-t)^2 + \frac{1}{2}\,t^2} \ = \ \frac{1}{\sqrt{2\,\pi}}\,\mathrm{e}^{\frac{1}{2}\,t^2}\ \mathrm{e}^{-\frac{1}{2}\,(x-t)^2}$$

- On effectue alors le changement de variable affine $\boxed{u = x - t}$.

  $u = x - t$ (et donc $x = u - t$)
  $\hookrightarrow du = dx$ et $dx = du$
  - $x = -\infty \ \Rightarrow\ u = -\infty$
  - $x = +\infty \ \Rightarrow\ u = +\infty$

  On obtient alors :
  $$\begin{aligned}
  \int_{-\infty}^{+\infty} \mathrm{e}^{t\,x}\ f_Z(x)\ dx &= \mathrm{e}^{\frac{1}{2}\,t^2} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\,\pi}}\ \mathrm{e}^{-\frac{1}{2}\,(x-t)^2}\ dx \\
  &= \mathrm{e}^{\frac{1}{2}\,t^2} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\,\pi}}\ \mathrm{e}^{-\frac{1}{2}\,u^2}\ du \\
  &= \mathrm{e}^{\frac{1}{2}\,t^2} \int_{-\infty}^{+\infty} f_Z(u)\ du \ = \ \mathrm{e}^{\frac{1}{2}\,t^2} \qquad \textit{(car } f_Z \textit{ est une densité de probabilité)}
  \end{aligned}$$

  $$\boxed{\forall t \in \mathbb{R},\ M_Z(t) = \exp\left(\frac{t^2}{2}\right)}$$

□

c) En déduire la valeur de tous les cumulants d'une variable aléatoire qui suit une loi normale d'espérance $\mu \in \mathbb{R}$ et d'écart-type $\sigma \in \mathbb{R}_+^*$.

Démonstration.

- On note Y la v.a.r. définie par : $Y = \sigma\,Z + \mu$ (où Z est toujours une v.a.r. de loi $\mathcal{N}(0,1)$).
  Alors : $Y \hookrightarrow \mathcal{N}\left(\mu, \sigma^2\right)$.
  De plus, d'après la question ***5.b)***, pour tout t tel que $\sigma\,t \in \mathcal{D}_Z$:
  $$K_Y(t) \ = \ \mu\,t + K_Z(\sigma\,t)$$
- Déterminons donc d'abord K_Z.
  × Tout d'abord, remarquons que pour tout $t \in \mathbb{R}$: $M_Z(t) = \mathrm{e}^{\frac{1}{2}\,t^2} > 0$.
  Ainsi, K_Z est définie sur $\mathbb{R}$, *i.e.* : $\mathcal{D}_Z = \mathbb{R}$.
  × De plus, pour tout $t \in \mathbb{R}$:
  $$K_Z(t) \ = \ \ln\left(\mathrm{e}^{\frac{1}{2}\,t^2}\right) \ = \ \frac{1}{2}\,t^2$$

- Soit $t \in \mathbb{R}$, alors $\sigma\, t \in \mathbb{R} = \mathcal{D}_Z$. Ainsi, d'après la question ***5.b)*** :

$$K_Y(t) \;=\; \mu\, t + K_Z(\sigma\, t) \;=\; \mu\, t + \frac{1}{2}\,(\sigma\, t)^2$$

$$\boxed{\text{Finalement : } K_Y : t \mapsto \mu\, t + \frac{\sigma^2}{2}\, t^2.}$$

- La fonction K_Y est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ en tant que fonction polynomiale.

$$\boxed{\text{Ainsi, la v.a.r. } Y \text{ admet des cumulants à tout ordre.}}$$

- On a alors, pour tout $t \in \mathbb{R}$:
  × $K_Y'(t) = \mu + \sigma^2\, t$ et ainsi : $Q_1(Y) = K_Y'(0) = \mu$.
  × $K_Y''(t) = \sigma^2$ et ainsi : $Q_2(Y) = K_Y''(0) = \sigma^2$.
  × pour tout $p \in [\![3, +\infty[\![$, $K_Y^{(p)}(t) = 0$ et ainsi : $Q_p(Y) = K_Y^{(p)}(0) = 0$.

$$\boxed{\text{En conclusion : } Q_1(Y) = \mu,\ Q_2(Y) = \sigma^2 \text{ et pour tout } p \in \mathbb{N}^* \setminus \{2\},\ Q_p(Y) = 0.}$$

□

11. Soit $(T_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telles que, pour tout $n \in \mathbb{N}^*$, la variable aléatoire T_n suit la loi de Poisson de paramètre n. Pour tout $n \in \mathbb{N}^*$, on pose : $W_n = \dfrac{T_n - n}{\sqrt{n}}$.

a) Justifier la convergence en loi de la suite de variables aléatoires $(W_n)_{n\in\mathbb{N}^*}$ vers une variable aléatoire W.

Démonstration.

- Pour tout $i \in \mathbb{N}^*$, notons X_i une v.a.r. de loi de Poisson $\mathcal{P}(1)$. On suppose de plus que les v.a.r. X_i sont indépendantes.
  Alors, par stabilité des lois de Poisson, la v.a.r. $S_n = \sum\limits_{i=1}^{n} X_i$ suit la loi de Poisson $\mathcal{P}(n)$.
  On en déduit que S_n et T_n suivent la même loi. Ainsi, les v.a.r. $\dfrac{S_n - n}{\sqrt{n}}$ et $\dfrac{T_n - n}{\sqrt{n}} = W_n$ ont même loi.
- La suite $(X_i)_{i\in\mathbb{N}^*}$ est constituée de v.a.r. :
  × indépendantes,
  × de même loi,
  × admettant la même espérance 1,
  × admettant la même variance $1 \neq 0$.
  Ainsi, comme $S_n \hookrightarrow \mathcal{P}(n)$, la v.a.r. centrée réduite associée à S_n est :

$$S_n^* \;=\; \frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathbb{V}(S_n)}} \;=\; \frac{S_n - n}{\sqrt{n}}$$

- Alors, d'après le théorème central limite :

$$S_n^* \xrightarrow[n\to+\infty]{\mathscr{L}} W \quad \text{où} \quad W \hookrightarrow \mathcal{N}(0,1)$$

  On a donc : $\forall t \in \mathbb{R},\ \lim\limits_{n\to+\infty} F_{S_n^*}(t) = \Phi(t)$.
  (la fonction Φ est la fonction de répartition de W)
- Or les v.a.r. $S_n^* = \dfrac{S_n - n}{\sqrt{n}}$ et $W_n = \dfrac{T_n - n}{\sqrt{n}}$ ont même loi, donc même fonction de répartition.
  On en déduit : $\forall x \in \mathbb{R},\ \lim\limits_{n\to+\infty} F_{W_n}(t) = \Phi(t)$.

$$\boxed{\text{On en conclut que la suite de v.a.r. } (W_n)_{n\in\mathbb{N}^*} \text{ converge en loi vers } W \text{ où } W \hookrightarrow \mathcal{N}(0,1).}$$

Commentaire

- La forme de la v.a.r. $W_n = \dfrac{T_n - n}{\sqrt{n}}$ dans le contexte de convergence de v.a.r. doit faire penser **systématiquement** au théorème central limite (TCL).
- De manière générale, si les X_i ont pour espérance m et variance σ^2, alors :

$$S_n^* \ = \ \frac{S_n - n\,m}{\sigma\,\sqrt{n}}$$

On voit bien apparaître la division par $\sqrt{n}$ qui est très caractéristique de l'utilisation du TCL.

- Rappelons aussi que si on pose : $V_n = \dfrac{1}{n} \sum\limits_{i=1}^{n} X_i$, alors :

$$V_n^* \ = \ \sqrt{n}\ \frac{V_n - m}{\sigma}$$

(on notera : $V_n^* = S_n^*$)
On voit alors apparaître le produit par $\sqrt{n}$ qui est aussi caractéristique du TCL.

- Pour pouvoir justifier de l'application de ce théorème dans cette question, on introduit la v.a.r. S_n qui suit la même loi que T_n et satisfait bien aux hypothèses du TCL.
On rappelle, comme remarqué en question ***3.c)***, que deux v.a.r. qui suivent la même loi ne sont pas forcément égales (on n'a d'ailleurs pas du tout besoin dans cette question de l'égalité $S_n = T_n$). Cependant on peut raisonnablement penser qu'un candidat précisant « $T_n = \sum\limits_{i=1}^{n} X_i$ » (et non : $S_n = \sum\limits_{i=1}^{n} X_i$) ne serait pas sanctionné.

□

b) Déterminer la fonction K_{W_n}.

Démonstration.

- Comme $T_n \hookrightarrow \mathcal{P}(n)$, d'après la question ***9.a)*** :

$$K_{T_n} : t \mapsto n(\mathrm{e}^t - 1)$$

- De plus : $W_n = \dfrac{T_n - n}{\sqrt{n}} \ = \ \dfrac{1}{\sqrt{n}}\, T_n - \dfrac{n}{\sqrt{n}} \ = \ \dfrac{1}{\sqrt{n}}\, T_n - \sqrt{n}$.
Ainsi, soit $t \in \mathbb{R}$, alors $\dfrac{1}{\sqrt{n}}\, t \in \mathbb{R}$ et, d'après la question ***5.b)*** :

$$K_{W_n}(t) \ = \ -\sqrt{n}\,t + K_{T_n}\left(\frac{1}{\sqrt{n}}\ t\right) \ = \ -\sqrt{n}\,t + n\left(\mathrm{e}^{\frac{1}{\sqrt{n}}\ t} - 1\right)$$

$$\boxed{\text{Finalement : } K_{W_n} : t \mapsto -\sqrt{n}\,t + n\left(\mathrm{e}^{\frac{t}{\sqrt{n}}} - 1\right).}$$

□

c) Montrer que pour tout $t \in \mathbb{R}$, on a : $\lim\limits_{n\to+\infty} K_{W_n}(t) = K_W(t)$.

Démonstration.
Soit $t \in \mathbb{R}$.

- Tout d'abord : $e^x = 1 + x + \dfrac{x^2}{2} + \underset{x\to 0}{o}(x^2)$.
  Or : $\lim\limits_{n\to+\infty} \dfrac{t}{\sqrt{n}} = 0$. On peut donc appliquer le développement limité précédent en choisissant $x = \dfrac{t}{\sqrt{n}}$. On obtient :

$$e^{\frac{t}{\sqrt{n}}} = 1 + \frac{t}{\sqrt{n}} + \frac{\left(\frac{t}{\sqrt{n}}\right)^2}{2} + \underset{n\to+\infty}{o}\left(\left(\frac{t}{\sqrt{n}}\right)^2\right) = 1 + \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right)$$

- On en déduit, d'après la question précédente :

$$\begin{aligned} K_{W_n}(t) &= -\sqrt{n}\,t + n\left(\cancel{1} + \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right) - \cancel{1}\right) \\ &= -\sqrt{n}\,t + n\,\frac{t}{\sqrt{n}} + \frac{t^2}{2} + \underset{n\to+\infty}{o}(1) \\ &= -\cancel{\sqrt{n}\,t} + \cancel{\sqrt{n}\,t} + \frac{t^2}{2} + \underset{n\to+\infty}{o}(1) \\ &= \frac{t^2}{2} + \underset{n\to+\infty}{o}(1) \end{aligned}$$

  On en déduit : $\lim\limits_{n\to+\infty} K_{W_n}(t) = \dfrac{t^2}{2}$.

- Or, comme $W \hookrightarrow \mathcal{N}(0,1)$, d'après la question ***10.c)*** : $K_W(t) = \dfrac{t^2}{2}$.

$$\boxed{\text{Finalement, pour tout } t \in \mathbb{R} : \lim_{n\to+\infty} K_{W_n}(t) = \frac{t^2}{2} = K_W(t).}$$

Commentaire

- Soit f une fonction et $(a, x_0) \in \mathbb{R}^2$.
  On rappelle la propriété utilisée dans le deuxième point :

$$\lim_{x\to x_0} f(x) = a \quad \Leftrightarrow \quad f(x) = a + \underset{x\to x_0}{o}(1)$$

- On peut remarquer que cette question ***11.*** nous fait démontrer, **dans le cas particulier d'une loi de Poisson**, l'implication suivante :

$$W_n \underset{n\to+\infty}{\overset{\mathscr{L}}{\longrightarrow}} W \quad \Rightarrow \quad \forall t \in \mathbb{R},\ K_{W_n}(t) \underset{n\to+\infty}{\longrightarrow} K_W(t)$$

  On peut se poser la question de la généralisation de cette propriété. En effet, il existe un lien entre M_X et F_X, mais cela serait hors de portée du programme ECE.

□

Partie III. Cumulant d'ordre 4

Dans cette partie, on considère une variable aléatoire X telle que M_X est de classe $\mathcal{C}^4$ sur un intervalle ouvert I contenant l'origine.
On admet alors que X possède des moments jusqu'à l'ordre 4 qui coïncident avec les dérivées successives de la fonction M_X en 0. Autrement dit, pour tout $k \in [\![1, 4]\!]$, on a : $M_X^{(k)}(0) = \mathbb{E}(X^k)$.
De plus, on pose : $\mu_4(X) = \mathbb{E}\left(\big(X - \mathbb{E}(X)\big)^4\right)$.

Commentaire

Dans la définition de μ_4, on sous-entend que la v.a.r. $\big(X - \mathbb{E}(X)\big)^4$ admet une espérance.
Dans le cours, on démontre les propriétés suivantes :

× si X admet une espérance, il en est de même de $X - \mathbb{E}(X)$ (c'est la v.a.r. centrée associée à X).
Dans ce cas, on a : $\mu_1(X) = \mathbb{E}\big(X - \mathbb{E}(X)\big) = 0$.

× si X admet un moment d'ordre 2, il en est de même de $X - \mathbb{E}(X)$.
Dans ce cas, on a : $\mu_2(X) = \mathbb{E}\Big(\big(X - \mathbb{E}(X)\big)^2\Big) = \mathbb{V}(X)$.
(la variance est le moment centré d'ordre 2)

De manière générale si X admet un moment d'ordre $n \in \mathbb{N}$, il en est de même de $X - \mathbb{E}(X)$.
En effet :

$$\big(X - \mathbb{E}(X)\big)^n = \sum_{k=0}^{n} \binom{n}{k} \big(-\mathbb{E}(X)\big)^{n-k} X^k$$

Ainsi, la v.a.r. $\big(X - \mathbb{E}(X)\big)^n$ admet une espérance en tant que combinaison linéaire des v.a.r. X^0, X_1, $\dots$, X^n, qui admettent toutes une espérance.

12. Justifier les égalités : $Q_1(X) = \mathbb{E}(X)$ et $Q_2(X) = \mathbb{V}(X)$.

Démonstration.

- La fonction K_X est de classe $\mathcal{C}^4$ sur I car elle est la composée $K_X = \ln \circ M_X$ où :
  × la fonction M_X est :
    – est de classe $\mathcal{C}^4$ sur I,
    – telle que $M_X(I) \subset \,]0, +\infty[$.
    En effet, comme : $\forall t \in I$, $\mathrm{e}^{tX} > 0$, on a : $\forall t \in I$, $M_X(t) = \mathbb{E}(\mathrm{e}^{tX}) > 0$.
  × la fonction ln est de classe $\mathcal{C}^4$ sur $]0, +\infty[$.
- Déterminons les dérivées successives de K_X. Soit $t \in I$.
  × Tout d'abord : $K_X'(t) = \dfrac{1}{M_X(t)} \times M_X'(t)$.

$$\begin{aligned}
Q_1(X) &= K_X'(0) && \textit{(par définition)} \\
&= \frac{1}{M_X(0)} \times M_X'(0) && \textit{(d'après le résultat précédent en } t = 0 \in I\textit{)} \\
&= \frac{1}{M_X(0)} \times \mathbb{E}(X) && \textit{(d'après le résultat admis dans l'énoncé)} \\
&= \frac{1}{\mathbb{E}(\mathrm{e}^0)}\, \mathbb{E}(X) \ = \ \mathbb{E}(X)
\end{aligned}$$

$$\boxed{Q_1(X) = \mathbb{E}(X)}$$

× Ensuite : $K_X''(t) = \left(-\dfrac{1}{\big(M_X(t)\big)^2} \times M_X'(t)\right)\ M_X'(t) + \dfrac{1}{M_X(t)} \times M_X''(t)$.

$$\begin{array}{rcll} Q_1(X) & = & K_X''(0) & \textit{(par définition)} \\ & = & \left(-\dfrac{1}{\big(M_X(0)\big)^2} \times M_X'(0)\right)\ M_X'(0) + \dfrac{1}{M_X(0)} \times M_X''(0) & \textit{(d'après le résultat précédent en } t = 0 \in I\textit{)} \\ & = & -\dfrac{1}{(1)^2} \times \mathbb{E}(X) \times \mathbb{E}(X) + \dfrac{1}{1} \times \mathbb{E}\big(X^2\big) & \textit{(d'après le résultat admis dans l'énoncé)} \\ & = & \mathbb{E}\big(X^2\big) - \big(\mathbb{E}(X)\big)^2 \ = \ \mathbb{V}(X) & \textit{(d'après la formule de Kœnig-Huygens)} \end{array}$$

$$\boxed{Q_2(X) = \mathbb{V}(X)}$$

Commentaire

- On se sert ici du résultat stipulant que toute v.a.r. X qui admet une espérance vérifie :

$$\boxed{X > 0 \quad \Rightarrow \quad \mathbb{E}(X) > 0} \ (*)$$

Ce résultat n'est pas officiellement au programme des classes préparatoires commerciales. Seul le résultat plus faible suivant (nommé parfois positivité de l'espérance) figure :

$$\boxed{X \geqslant 0 \quad \Rightarrow \quad \mathbb{E}(X) \geqslant 0} \ (**)$$

On notera au passage que l'on obtient le même résultat en supposant l'hypothèse plus faible : $\mathbb{P}\big([X \geqslant 0]\big) = 1$.

- Pour démontrer le résultat $(*)$, on peut procéder par l'absurde.
Supposons $X > 0$ et $\texttt{NON}(\mathbb{E}(X) > 0)$ (autrement dit $\mathbb{E}(X) \leqslant 0$).
  × Comme $X > 0 \geqslant 0$, alors, d'après $(**)$: $\mathbb{E}(X) \geqslant 0$.
  Comme $\mathbb{E}(X) \leqslant 0$, on en déduit $\mathbb{E}(X) = 0$.
  × D'après l'inégalité de Markov : $\forall a > 0,\ \mathbb{P}\big([X > a]\big) \leqslant \dfrac{\mathbb{E}(X)}{a} = 0$.
  Une probabilité étant toujours positive, on démontre ainsi : $\forall a > 0,\ \mathbb{P}\big([X > a]\big) = 0$.
  × Démontrons, à l'aide de cette dernière égalité : $\mathbb{P}\big([X > 0]\big) = 0$.
  Pour ce faire, on remarque tout d'abord : $[X > 0] = \bigcup\limits_{n=1}^{+\infty} \left[X > \frac{1}{n}\right]$ *(démonstration par double inclusion)*
  On en déduit :

$$\begin{array}{rcll} \mathbb{P}\big([X > 0]\big) & = & \mathbb{P}\Big(\bigcup\limits_{n=1}^{+\infty} \left[X > \frac{1}{n}\right]\Big) & \\ & = & \lim\limits_{N \to +\infty} \mathbb{P}\Big(\bigcup\limits_{n=1}^{N} \left[X > \frac{1}{n}\right]\Big) & \textit{(d'après la propriété de la limite monotone)} \\ & = & \lim\limits_{N \to +\infty} \mathbb{P}\Big(\left[X > \frac{1}{N}\right]\Big) & \textit{(car } \Big(\left[X > \frac{1}{N}\right]\Big)_{N \in \mathbb{N}^*} \textit{ est une suite croissante d'événements)} \\ & = & \lim\limits_{N \to +\infty} 0 \ = \ 0 & \textit{(car } \mathbb{P}\Big(\left[X > \frac{1}{N}\right]\Big) = 0\textit{)} \end{array}$$

  Absurde ! □

13. Soit X_1 et X_2 deux variables aléatoires indépendantes et de même loi que X. On pose : $S = X_1 - X_2$.

a) Montrer que la variable aléatoire S possède un moment d'ordre 4 et établir l'égalité :

$$\mathbb{E}(S^4) \;=\; 2\,\mu_4(X) + 6\left(\mathbb{V}(X)\right)^2$$

Démonstration.

- D'après l'énoncé, les v.a.r. X_1 et X_2 ont même loi que X, v.a.r. qui admet une espérance. Ainsi $\mathbb{E}(X_1) = \mathbb{E}(X) = \mathbb{E}(X_2)$ et :

$$\begin{aligned} S &= X_1 - X_2 \\ &= X_1 - X_2 - \mathbb{E}(X_1) + \mathbb{E}(X_2) \\ &= \big(X_1 - \mathbb{E}(X_1)\big) - \big(X_2 - \mathbb{E}(X_2)\big) \end{aligned}$$

- On en déduit, à l'aide de la formule du binôme de Newton :

$$\begin{aligned} S^4 &= \Big(\big(X_1 - \mathbb{E}(X_1)\big) - \big(X_2 - \mathbb{E}(X_2)\big)\Big) \\ &= \sum_{k=0}^{4} \binom{4}{k} \left(X_1 - \mathbb{E}(X_1)\right)^{4-k} \Big(-\left(X_2 - \mathbb{E}(X_2)\right)^{k}\Big) \\ &= \sum_{k=0}^{4} \binom{4}{k} (-1)^k \left(X_1 - \mathbb{E}(X_1)\right)^{4-k} \left(X_2 - \mathbb{E}(X_2)\right)^{k} \end{aligned}$$

  Remarquons alors que :

  × les v.a.r. $X_1 - \mathbb{E}(X_1)$ et $X_2 - \mathbb{E}(X_2)$ admettent des moments à tout ordre $r \in [\![0,4]\!]$ car suivent la même loi que $X - \mathbb{E}(X)$ qui admet un moment d'ordre 4.

  × d'après le lemme des coalitions, pour tout entier $k \in [\![0,4]\!]$, les v.a.r. $\left(X_1 - \mathbb{E}(X_1)\right)^{4-k}$ et $\left(X_2 - \mathbb{E}(X_2)\right)^{k}$ sont indépendantes car X_1 et X_2 le sont.

  On en déduit que pour tout $k \in [\![0,4]\!]$, la v.a.r. $\left(X_1 - \mathbb{E}(X_1)\right)^{4-k} \left(X_2 - \mathbb{E}(X_2)\right)^{k}$ admet une espérance. Et par propriété de l'espérance :

$$\mathbb{E}\Big(\left(X_1 - \mathbb{E}(X_1)\right)^{4-k} \left(X_2 - \mathbb{E}(X_2)\right)^{k}\Big) \;=\; \mathbb{E}\Big(\left(X_1 - \mathbb{E}(X_1)\right)^{4-k}\Big)\; \mathbb{E}\Big(\left(X_2 - \mathbb{E}(X_2)\right)^{k}\Big)$$

$$\boxed{\text{La v.a.r. } S^4 \text{ admet une espérance comme combinaison linéaire de v.a.r. qui admettent une espérance.}}$$

- De plus, par linéarité de l'espérance :

$$\begin{aligned} \mathbb{E}(S^4) &= \sum_{k=0}^{4} \binom{4}{k} (-1)^k\, \mathbb{E}\Big(\left(X_1 - \mathbb{E}(X_1)\right)^{4-k}\Big)\, \mathbb{E}\Big(\left(X_2 - \mathbb{E}(X_2)\right)^{k}\Big) \\ &= \quad 1 \;\; \mathbb{E}\Big(\left(X_1 - \mathbb{E}(X_1)\right)^{4}\Big) \\ &\quad -4 \;\; \mathbb{E}\Big(\left(X_1 - \mathbb{E}(X_1)\right)^{3}\Big)\, \cancel{\mathbb{E}\Big(\left(X_2 - \mathbb{E}(X_2)\right)^{1}\Big)} && \textit{(car } X_2 - \mathbb{E}(X_2) \textit{ est la v.a.r. centrée associée à } X_2\textit{)} \\ &\quad +6 \;\; \mathbb{E}\Big(\left(X_1 - \mathbb{E}(X_1)\right)^{2}\Big)\, \mathbb{E}\Big(\left(X_2 - \mathbb{E}(X_2)\right)^{2}\Big) \\ &\quad -4 \;\; \cancel{\mathbb{E}\Big(\left(X_1 - \mathbb{E}(X_1)\right)^{1}\Big)}\, \mathbb{E}\Big(\left(X_2 - \mathbb{E}(X_2)\right)^{3}\Big) && \textit{(car } X_1 - \mathbb{E}(X_1) \textit{ est la v.a.r. centrée associée à } X_1\textit{)} \\ &\quad +1 \;\; \mathbb{E}\Big(\left(X_2 - \mathbb{E}(X_2)\right)^{4}\Big) \\ &= \mu_4(X_1) \;+\; 6\,\mathbb{V}(X_1)\;\mathbb{V}(X_2) \;+\; \mu_4(X_2) \end{aligned}$$

Enfin, comme X_1 et X_2 ont même loi que X :

$$\mathbb{V}(X_1) \;=\; \mathbb{V}(X) \;=\; \mathbb{V}(X_2) \quad \text{et} \quad \mu_4(X_1) \;=\; \mu_4(X) \;=\; \mu_4(X_2)$$

$$\boxed{\text{Et ainsi : } \mathbb{E}\big(S^4\big) \;=\; 2\,\mu_4(X_1) + 6\,\big(\mathbb{V}(X)\big)^2.}$$

□

b) Montrer que les fonctions M_S et K_S sont de classe $\mathcal{C}^4$ sur I et que pour tout $t \in I$, on a :

$$M_S^{(4)}(t) \;=\; K_S^{(4)}(t)\,M_S(t) + 3\,K_S^{(3)}(t)\,M_S'(t) + 3\,K_S''(t)\,M_S''(t) + K_S'(t)\,M_S^{(3)}(t)$$

Démonstration.
Soit $t \in I$ tel que $-t \in I$.

- Comme $-t \in I$, la quantité $M_X(-t)$ est bien définie. On a alors :

$$\begin{array}{ccccccc} M_X(-t) & = & \mathbb{E}\big(e^{-t\,X}\big) & = & \mathbb{E}\big(e^{t\,(-X)}\big) & = & M_{-X}(t) \\ \shortparallel & & & & & & \shortparallel \\ M_{X_2}(-t) & & & & & & M_{-X_2}(t) \end{array}$$

(car X et X_2 ont même loi)

$$\boxed{\text{Ainsi, pour tout } t \in I \text{ tel que } -t \in I, \text{ on a : } M_{-X_2}(t) = M_{X_2}(-t).}$$

- D'autre part :

$$e^{t\,S} \;=\; e^{t\,(X_1 - X_2)} \;=\; e^{t\,X_1} \times e^{-t\,X_2}$$

Les v.a.r. $e^{t\,X_1}$ et $e^{-t\,X_2}$:

× admettent toutes les deux une espérance car M_{X_1} est définie en t et M_{X_2} est définie en $-t$ (car on a supposé $-t \in I$).

× sont indépendantes d'après le lemme des coalitions puisque X_1 et X_2 le sont.

On en déduit que la v.a.r. $e^{t\,X_1} \times e^{-t\,X_2}$ admet une espérance, donnée par :

$$\begin{array}{ccccc} \mathbb{E}\big(e^{t\,S}\big) & = & \mathbb{E}\big(e^{t\,X_1} \times e^{-t\,X_2}\big) & = & \mathbb{E}\big(e^{t\,X_1}\big) \times \mathbb{E}\big(e^{-t\,X_2}\big) \\ \shortparallel & & & & \shortparallel \\ M_S(t) & & & & M_{X_1}(t) \times M_{X_2}(-t) \\ & & & & \shortparallel \\ & & & & M_X(t) \times M_X(-t) \end{array}$$

(d'après le point précédent)

(car les v.a.r. X_2 et X_1 ont même loi que X)

$$\boxed{\begin{array}{c}\text{Ainsi, pour tout intervalle } J \subset I \text{ symétrique par rapport à l'origine, on a :} \\ \forall t \in J,\ M_S(t) \;=\; M_X(t) \times M_X(-t).\end{array}}$$

> En particulier, la fonction M_S est de classe $\mathcal{C}^4$ sur tout intervalle $J \subset I$ symétrique par rapport à l'origine, comme produit de fonctions de classe $\mathcal{C}^4$ sur cet intervalle. On en déduit, comme en question ***12***, que la fonction K_S est elle aussi de classe $\mathcal{C}^4$ sur J.

Dans la suite, on note $J \subset I$ un intervalle symétrique par rapport à l'origine. Soit $t \in J$.

- Tout d'abord : $K_S'(t) \;=\; \dfrac{1}{M_S(t)} \times M_S'(t)$.

$$\boxed{\text{On en déduit : } \forall t \in J,\ M_S'(t) \;=\; K_S'(t) \times M_S(t).}$$

- En remarquant : $\big(M_S\big)^{(4)}(t) = \big(M'_S\big)^{(3)}(t)$, on obtient :

$$\begin{aligned}
& \big(M_S\big)^{(4)}(t) \\
= \ & \big(K'_S \times M_S\big)^{(3)}(t) \\
= \ & \Big(\big(K'_S \times M_S\big)'\Big)^{(2)}(t) \\
= \ & \Big(K''_S \times M_S + K'_S \times M'_S\Big)^{(2)}(t) \\
= \ & \Big(\big(K''_S \times M_S + K'_S \times M'_S\big)'\Big)^{(1)}(t) \\
= \ & \Big(\big(K'''_S \times M_S + K''_S \times M'_S\big) + \big(K''_S \times M'_S + K'_S \times M''_S\big)\Big)^{(1)}(t) \\
= \ & \Big(K'''_S \times M_S + 2\,K''_S \times M'_S + K'_S \times M''_S\Big)^{(1)}(t) \\
= \ & \Big(\big(K''''_S \times M_S + K'''_S \times M'_S\big) + 2\,\big(K'''_S \times M'_S + K''_S \times M''_S\big) + \big(K''_S \times M''_S + K'_S \times M'''_S\big)\Big)(t) \\
= \ & \Big(K''''_S \times M_S + 3\,K'''_S \times M'_S + 3\,K''_S \times M''_S + K'_S \times M'''_S\Big)(t) \\
= \ & K_S^{(4)}(t)\, M_S(t) + 3\,K_S^{(3)}(t)\, M'_S(t) + 3\,K''_S(t)\, M''_S(t) + K'_S(t)\, M_S^{(3)}(t)
\end{aligned}$$

$$\boxed{\forall t \in J,\ M_S^{(4)}(t) = K_S^{(4)}(t)\, M_S(t) + 3\,K_S^{(3)}(t)\, M'_S(t) + 3\,K''_S(t)\, M''_S(t) + K'_S(t)\, M_S^{(3)}(t)}$$

Commentaire

- Lors de l'étude de $M_S(t)$ (pour $t \in I$) la quantité $M_{X_2}(-t)$ apparaît naturellement. Or, cette quantité existe seulement si $-t \in I$. C'est pourquoi on a décidé dans cette question de restreindre à la démonstration à un intervalle $J \subset I$ symétrique par rapport à l'origine.
- Dans la correction, on a déterminé la dérivée quatrième de M_S en dérivant successivement trois fois la fonction M'_S. On aurait pu utiliser directement la formule de Leibniz qui stipule que si $f, g : I \to \mathbb{R}$ sont deux fonctions n fois dérivables sur un intervalle I, on a :

$$\boxed{\big(f \times g\big)^{(n)} \ = \ \sum_{k=0}^{n} \binom{n}{k}\ f^{(n-k)} \times g^{(k)}}$$

Cette formule n'apparaît pas explicitement dans le programme ECE (mais est bien présente dans le programme ECS de première année). Il est toutefois très classique de la présenter lors de la première année ECE. On pouvait l'utiliser directement ici et écrire :

$$\begin{aligned}
\big(M_S\big)^{(4)}(t) \ &= \ \big(M'_S\big)^{(3)}(t) \ = \ \sum_{k=0}^{3} \binom{3}{k}\ {K'_S}^{(3-k)} \times \big(M_S\big)^{(k)} \\
&= \ \sum_{k=0}^{3} \binom{3}{k}\ K_S^{(4-k)} \times \big(M_S\big)^{(k)} \\
&= \ K_S^{(4)}(t)\, M_S(t) + 3\,K_S^{(3)}(t)\, M'_S(t) + 3\,K''_S(t)\, M''_S(t) + K'_S(t)\, M_S^{(3)}(t)
\end{aligned}$$

La présence des coefficients 1, 3, 3, 1 dans la formule à démontrer doit mettre sur la piste de l'utilisation d'une formule utilisant les coefficients binomiaux. □

c) En déduire l'égalité : $\mathbb{E}(S^4) \ = \ Q_4(S) + 3\left(\mathbb{V}(S)\right)^2$.

Démonstration.

- Notons tout d'abord que la fonction M_S est de classe $\mathcal{C}^4$ sur l'intervalle J (qu'on peut choisir ouvert) qui contient l'origine. On admet donc, comme dans l'énoncé :
$$\forall k \in [\![1,4]\!],\ M_S^{(k)}(0) = \mathbb{E}\big(S^k\big)$$
- La v.a.r. S vérifiant les mêmes hypothèses que la v.a.r. X de début de la **Partie III**, on en déduit, que le résultat de la question ***12*** est vérifié pour la v.a.r. S. Plus précisément :
$$Q_1(S) = \mathbb{E}(S) \quad \text{et} \quad Q_2(S) = \mathbb{V}(S)$$
Remarquons au passage :
$$\begin{array}{rcll} \mathbb{E}(S) & = & \mathbb{E}(X_1 - X_2) & \\ & = & \mathbb{E}(X_1) - \mathbb{E}(X_2) & \textit{(par linéarité de l'espérance)} \\ & = & \mathbb{E}(X) - \mathbb{E}(X) \ = \ 0 & \textit{(car } X_1 \textit{ et } X_2 \textit{ ont même loi que } X\textit{)} \end{array}$$
Et ainsi :
$$\begin{array}{rcll} \mathbb{V}(S) & = & \mathbb{E}\big(S^2\big) - \cancel{\big(\mathbb{E}(S)\big)^2} & \textit{(d'après la formule de Kœnig-Huygens)} \\ & = & \mathbb{E}\big(S^2\big) & \end{array}$$
- On applique alors l'égalité précédente à $t = 0 \in J$:
$$\begin{array}{rl} & M_S^{(4)}(0) \\ = & K_S^{(4)}(0) \times M_S(0) \ + \ 3\,K_S^{(3)}(0) \times M_S'(0) \ + \ 3\,K_S''(0) \times M_S''(0) \ + \ K_S'(0) \times M_S^{(3)}(0) \\ = & Q_4(S) \times M_S(0) \ + \ 3\,Q_3(S) \times M_S'(0) \ + \ 3\,Q_2(S) \times M_S''(0) \ + \ Q_1(S) \times M_S^{(3)}(0) \\ = & Q_4(S) \times \mathbb{E}\big(\mathrm{e}^0\big) \ + \ 3\,Q_3(S) \times \mathbb{E}(S) \ + \ 3\,Q_2(S) \times \mathbb{E}\big(S^2\big) \ + \ Q_1(S) \times \mathbb{E}\big(S^3\big) \\ = & Q_4(S) \ + \ 3\,Q_3(S) \times \cancel{\mathbb{E}(S)} \ + \ 3\,\mathbb{V}(S) \times \mathbb{E}\big(S^2\big) \ + \ \cancel{\mathbb{E}(S)} \times \mathbb{E}\big(S^3\big) \\ = & Q_4(S) \ + \ 3\,\mathbb{V}(S) \times \mathbb{V}(S) \end{array}$$
$$\boxed{\text{On a bien : } \mathbb{E}(S^4) \ = \ M_S^{(4)}(0) \ = \ Q_4(S) + 3\left(\mathbb{V}(S)\right)^2.}$$ □

14. Justifier que le cumulant d'ordre 4 de X est donné par la relation : $Q_4(X) \ = \ \mu_4(X) - 3\left(\mathbb{V}(X)\right)^2$.

Démonstration.

- Tout d'abord :
$$\begin{array}{rcll} Q_4(S) & = & \mathbb{E}(S^4) - 3\left(\mathbb{V}(S)\right)^2 & \textit{(d'après la question } \textbf{13.c)}\textit{)} \\ & = & \left(2\,\mu_4(X) + 6\left(\mathbb{V}(X)\right)^2\right) - 3\left(\mathbb{V}(S)\right)^2 & \textit{(d'après la question } \textbf{13.a)}\textit{)} \\ & = & \left(2\,\mu_4(X) + 6\left(\mathbb{V}(X)\right)^2\right) - 3\left(\mathbb{V}(X_1 - X_2)\right)^2 & \\ & = & \left(2\,\mu_4(X) + 6\left(\mathbb{V}(X)\right)^2\right) - 3\left(\mathbb{V}(X_1) + \mathbb{V}(-X_2)\right)^2 & \textit{(car } X_1 \textit{ et } -X_2 \textit{ sont indépendantes d'après le lemme des coalitions)} \\ & = & \left(2\,\mu_4(X) + 6\left(\mathbb{V}(X)\right)^2\right) - 3\left(\mathbb{V}(X_1) + (-1)^2\,\mathbb{V}(X_2)\right)^2 & \\ & = & \left(2\,\mu_4(X) + 6\left(\mathbb{V}(X)\right)^2\right) - 3\left(\mathbb{V}(X) + \mathbb{V}(X)\right)^2 & \textit{(car } X_1 \textit{ et } X_2 \textit{ ont même loi que } X\textit{)} \\ & = & 2\,\mu_4(X) + 6\left(\mathbb{V}(X)\right)^2 - 12\left(\mathbb{V}(X)\right)^2 \ = \ 2\,\mu_4(X) - 6\left(\mathbb{V}(X)\right)^2 & \end{array}$$

- Il reste alors à exprimer $Q_4(S)$ en fonction de $Q_4(X)$.
  Rappelons tout d'abord qu'on a démontré en question ***13.b)*** que pour tout $t \in J$:
  $$M_S(t) \;=\; M_X(t)\; M_X(-t)$$
  On en conclut :
  $$\begin{array}{rcl} K_S(t) & = & \ln\big(M_S(t)\big) \\ & = & \ln\big(M_X(t) \times M_X(-t)\big) \\ & = & \ln\big(M_X(t)\big) + \ln\big(M_X(-t)\big) \\ & = & K_X(t) + K_X(-t) \end{array}$$
  Ainsi, par dérivations successives des deux membres de cette égalité :
  $$\begin{array}{lrcl} & K'_S(t) & = & K'_X(t) - K'_X(-t) \\ \text{donc} & K''_S(t) & = & K''_X(t) - \big(-K''_X(-t)\big) \\ & & = & K'_X(t) + K'_X(-t) \\ \text{et} & K_S^{(3)}(t) & = & K_X^{(3)}(t) - K_X^{(3)}(-t) \\ \text{enfin} & K_S^{(4)}(t) & = & K_X^{(4)}(t) + K_X^{(4)}(-t) \end{array}$$
  En particulier, pour $t = 0 \in J$:
  $$\begin{array}{ccccc} K_S^{(4)}(0) & = & K_X^{(4)}(0) + K_X^{(4)}(0) & = & 2\,K_X^{(4)}(0) \\ \shortparallel & & & & \shortparallel \\ Q_4(S) & & & & 2\,Q_4(X) \end{array}$$
- On en conclut, d'après ce qui précède :
  $$2\,Q_4(X) \;=\; 2\,\mu_4(X) - 6\,\big(\mathbb{V}(X)\big)^2$$
  $$\boxed{\text{Finalement : } Q_4(X) \;=\; \mu_4(X) - 3\,\big(\mathbb{V}(X)\big)^2.}$$
  □

HEC 2020 - couple de variables aléatoires à densité, loi exponentielle, loi de Bernoulli, loi binomiale, inégalité de Boole

On s'intéresse dans ce sujet au problème de la *double dépense* de *bitcoins* par un groupe d'individus mal intentionnés.
On rappelle que le bitcoin est une monnaie virtuelle dont l'utilisation pour des transactions est associée à une structure unique appelée *blockchain*, partagée sur le réseau des usagers de cette monnaie et ayant pour but de sécuriser ces transactions.
La modélisation étudiée ne nécessite pas de connaissances particulières sur le *bitcoin* et la *blockchain*.

Partie I - Deux résultats généraux

On démontre dans cette partie deux résultats préliminaires, aux questions ***5.*** et ***6***. Ces résultats seront utilisés dans la suite du sujet et pourront être admis.

Calcul d'une probabilité

Soient X et Y deux variables aléatoires sur un espace probabilisé, à densité et indépendantes.
On note F_X et F_Y les fonctions de répartition de X et Y.
On suppose que Y est à valeurs positives et possède une densité f_Y dont la restriction à $[0, +\infty[$ est continue sur cet intervalle.
Pour tout $x \in \mathbb{R}_+$, on pose : $H(x) = \mathbb{P}(\, [X \leqslant Y] \cap [Y \leqslant x]\,)$.

1. a) Montrer que H est une fonction croissante sur $\mathbb{R}_+$ qui admet une limite finie en $+\infty$.

Démonstration.

- Soit $(x, y) \in \mathbb{R}_+ \times \mathbb{R}_+$.

$$\begin{array}{lrcl} \text{On suppose} & x & \leqslant & y \\ \text{on a alors} & [Y \leqslant x] & \subset & [Y \leqslant y] \\ \text{donc} & [X \leqslant Y] \cap [Y \leqslant x] & \subset & [X \leqslant Y] \cap [Y \leqslant y] \\ \text{donc} & \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant x]\big) & \leqslant & \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant y]\big) \\ & \shortparallel & & \shortparallel \\ & H(x) & & H(y) \end{array}$$

$$\boxed{\text{On en conclut que la fonction } H \text{ est croissante sur } \mathbb{R}_+.}$$

Commentaire

- La seule difficulté de cette question est de connaître la définition de croissance d'une fonction. Pour les fonctions dérivables, la propriété de croissance est souvent obtenue à l'aide de la caractérisation à l'aide du signe de la dérivée. Rappelons cependant que la définition de croissance n'utilise pas de propriété de régularité de la fonction.
- Démontrons formellement l'inclusion : $[Y \leqslant x] \ \subset \ [Y \leqslant y]$.
  Soit $\omega \in \Omega$. Supposons $\omega \in [Y \leqslant x]$.
  On a alors $Y(\omega) \leqslant x$ et ainsi :
  $$Y(\omega) \leqslant x \leqslant y$$
  On en conclut : $\omega \in [Y \leqslant y]$.

- La fonciton H est :
  - $\times$ croissante sur $\mathbb{R}_+$.
  - $\times$ majorée. En effet, pour tout $x \in \mathbb{R}$, $\mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant x]\big) \leqslant 1$.

On en déduit, par le théorème de la limite monotone, que la fonction admet une limite finie en $+\infty$.

□

b) En utilisant la suite $\big(H(n)\big)_{n\in\mathbb{N}}$, démontrer : $\lim\limits_{x\to+\infty} H(x) = \mathbb{P}([X \leqslant Y])$.
Que vaut $H(0)$?

Démonstration.
On note $(\Omega, \mathscr{A}, \mathbb{P})$ l'espace probabilisé évoqué dans l'énoncé.

- Démontrons tout d'abord : $\Omega = \bigcup\limits_{k=0}^{+\infty} [Y \leqslant k]$. On procède par double inclusion.

($\subset$) Soit $\omega \in \Omega$. Notons $m = \lceil Y(\omega) \rceil$. On a alors :

$$Y(\omega) \leqslant \lceil Y(\omega) \rceil = m$$

Ainsi : $\omega \in [Y \leqslant m] \ \subset \ \bigcup\limits_{k=0}^{+\infty} [Y \leqslant k]$.

$$\boxed{\Omega \ \subset \ \bigcup_{k=0}^{+\infty} [Y \leqslant k]}$$

($\supset$) Comme pour tout $k \in \mathbb{N}$, $[Y \leqslant k] \in \mathscr{A}$, alors : $\bigcup\limits_{k=0}^{+\infty} [Y \leqslant k] \in \mathscr{A}$.

$$\boxed{\text{En particulier : } \bigcup_{k=0}^{+\infty} [Y \leqslant k] \ \subset \ \Omega.}$$

- On en déduit alors :

$$\begin{array}{rcl} [X \leqslant Y] & = & [X \leqslant Y] \cap \left(\bigcup\limits_{k=0}^{+\infty} [Y \leqslant k]\right) \\ & = & \bigcup\limits_{k=0}^{+\infty} \Big([X \leqslant Y] \cap [Y \leqslant k]\Big) \end{array}$$

$$\begin{array}{rcll} \text{Finalement } \ \mathbb{P}\big([X \leqslant Y]\big) & = & \mathbb{P}\left(\bigcup\limits_{k=0}^{+\infty} \Big([X \leqslant Y] \cap [Y \leqslant k]\Big)\right) & \\ & = & \lim\limits_{N\to+\infty} \mathbb{P}\left(\bigcup\limits_{k=0}^{N} \Big([X \leqslant Y] \cap [Y \leqslant k]\Big)\right) & \textit{(d'après le théorème de la limite monotone)} \\ & = & \lim\limits_{N\to+\infty} \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant N]\big) & \textit{(car } \big([X \leqslant Y] \cap [Y \leqslant k]\big)_{k\in\mathbb{N}} \textit{ est une suite croissante d'événements)} \\ & = & \lim\limits_{N\to+\infty} H(N) & \\ & = & \lim\limits_{x\to+\infty} H(x) & \end{array}$$

$$\boxed{\text{On a bien : } \lim_{x\to+\infty} H(x) \ = \ \mathbb{P}\big([X \leqslant Y]\big).}$$

Commentaire

- Cette question est à juger comme difficile car elle exige beaucoup d'initiatives de la part du candidat. Il y a là un saut de difficulté par rapport à la question précédente.
- On n'a pas détaillé ci-dessus le fait que $\big([X \leqslant Y] \cap [Y \leqslant k]\big)_{k\in\mathbb{N}}$ est une suite croissante d'événements. C'est une application directe de la question précédente. En effet, on a démontré que pour tout $(x,y) \in \mathbb{R}_+ \times \mathbb{R}_+$:

$$x \leqslant y \ \Rightarrow\ [X \leqslant Y] \cap [Y \leqslant x] \subset [X \leqslant Y] \cap [Y \leqslant y]$$

Soit $n \in \mathbb{N}$. En choisissant $x = n$ et $y = n+1$ on obtient le résultat souhaité, à savoir :

$$[X \leqslant Y] \cap [Y \leqslant n] \subset [X \leqslant Y] \cap [Y \leqslant n+1]$$

- Par définition :

$$\begin{array}{rcll} H(0) & = & \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant 0]\big) & \\ & \leqslant & \mathbb{P}\big([Y \leqslant 0]\big) & \textit{(car } [X \leqslant Y] \cap [Y \leqslant 0] \subset [Y \leqslant 0]\textit{)} \\ & = & \mathbb{P}\big([Y < 0]\big) & \textit{(car } Y \textit{ est une v.a.r. à densité)} \\ & = & \mathbb{P}(\varnothing) & \textit{(car } Y \textit{ est à valeurs positives)} \\ & = & 0 & \end{array}$$

$$\boxed{H(0) = 0}$$

□

2. Soit (u,v) un couple de réels positifs tels que : $u < v$.

a) Montrer : $H(v) - H(u) = \mathbb{P}(\,[X \leqslant Y] \cap [u < Y \leqslant v]\,)$. Puis :

$$F_X(u)\ \frac{F_Y(v) - F_Y(u)}{v-u} \ \leqslant\ \frac{H(v) - H(u)}{v-u} \ \leqslant\ F_X(v)\ \frac{F_Y(v) - F_Y(u)}{v-u}$$

Démonstration.

- Il s'agit de démontrer :

$$\begin{array}{rcl} H(v) & = & H(u) + \mathbb{P}\big([X \leqslant Y] \cap [u < Y \leqslant v]\big) \\ & = & \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant u]\big) + \mathbb{P}\big([X \leqslant Y] \cap [u < Y \leqslant v]\big) \end{array}$$

où $H(v) = \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant v]\big)$. Pour ce faire, démontrons :

$$[X \leqslant Y] \cap [Y \leqslant v] \ = \ [X \leqslant Y] \cap [Y \leqslant u] \ \cup \ [X \leqslant Y] \cap [u < Y \leqslant v]$$

ou plus simplement : $[Y \leqslant v] \ = \ [Y \leqslant u] \ \cup \ [u < Y \leqslant v]$.

Commentaire

L'énoncé demande ici de démontrer une égalité entre probabilités de différents événements. Il est classique, pour ce faire, d'agir comme suit :

1) on démontre tout d'abord une égalité entre les événements concernés.

2) on applique alors l'application probabilité $\mathbb{P}$ de part et d'autre de l'égalité. On conclut à l'aide des propriétés de $\mathbb{P}$.

Il est à noter que l'égalité entre événements à démontrer est évidemment issue de l'égalité entre probabilité à démontrer. Pour ce faire, on aura en tête les triptyques :

union / incompatibilité / somme	intersection / indépendance / produit

On procède par double inclusion. Soit $\omega \in \Omega$.

$(\subset)$ Supposons $\omega \in [Y \leqslant v]$. Ainsi : $Y(\omega) \leqslant v$.
Deux cas se présentent alors :

× si $Y(\omega) \leqslant u$ alors $\omega \in [Y \leqslant u]$.

× si NON$(Y(\omega) \leqslant u)$ alors $Y(\omega) > u$.
Comme on sait de plus : $Y(\omega) \leqslant v$, on en conclut : $\omega \in [u < Y \leqslant v]$.

$$\boxed{\text{Finalement, on a bien : } \omega \in [Y \leqslant u] \cup [u < Y \leqslant v].}$$

$(\subset)$ Supposons $\omega \in [Y \leqslant u] \cup [u < Y \leqslant v]$. Ainsi : $Y(\omega) \leqslant u$ OU $u < Y(\omega) \leqslant v$.
Deux cas se présentent alors :

× si $Y(\omega) \leqslant u$ alors, comme $u < v$, on a $Y(\omega) \leqslant u < v$ et ainsi $\omega \in [Y \leqslant v]$.

× si NON$(Y(\omega) \leqslant u)$ alors, comme $Y(\omega) \leqslant u$ OU $u < Y(\omega) \leqslant v$, on a forcément : $u < Y(\omega) \leqslant v$. En particulier : $Y(\omega) \leqslant v$, et donc : $\omega \in [Y \leqslant v]$.

$$\boxed{\text{Finalement, on a bien : } \omega \in [Y \leqslant v].}$$

$$\boxed{\text{On en conclut : } [Y \leqslant v] \ = \ [Y \leqslant u] \ \cup \ [u < Y \leqslant v].}$$

Commentaire

- La démonstration de cette égalité a été développée ici afin d'illustrer la méthode. Cependant, cette égalité n'étant pas mentionnée dans l'énoncé, il est probable que l'écrire suffise à récupérer une grande partie des points alloués à la question.
- L'égalité initiale entre probabilités fait apparaître une différence entre probabilités de certains événements. Une telle égalité est généralement la conséquence d'une égalité entre événements où apparaît une différence ensembliste d'événements. Plus précisément, on pourrait mettre ici en place le raisonnement suivant :

$$[Y \leqslant v] \setminus [Y \leqslant u] = [u < Y \leqslant v] \quad \Rightarrow \quad \mathbb{P}\big([Y \leqslant v]\big) - \mathbb{P}\big([Y \leqslant u]\big) = \mathbb{P}\big([u < Y \leqslant v]\big)$$

Profitons-en pour rappeler que pour tout événement $(A, B) \in \mathscr{A} \times \mathscr{A}$, on a :

$$\mathbb{P}(A \setminus B) \ = \ \mathbb{P}\big(A \setminus A \cap B\big) \ = \ \mathbb{P}(A) - \mathbb{P}(A \cap B)$$

Afin de faciliter la résolution de cette question, on a préféré ici réordonner les termes de l'égalité de sorte à faire apparaître une somme entre probabilités d'événements. Une telle égalité est issue d'une réunion d'événements (le plus souvent incompatibles ou à tout le moins d'intersection négligeable) ce qui permet d'éviter d'avoir à gérer une différence ensembliste. Ce qui amène ici au raisonnement suivant :

$$[Y \leqslant u] \cup [u < Y \leqslant v] = [Y \leqslant v] \quad \Rightarrow \quad \mathbb{P}\big([Y \leqslant u]\big) + \mathbb{P}\big([u < Y \leqslant v]\big) = \mathbb{P}\big([Y \leqslant v]\big)$$

- Ainsi $\qquad [Y \leqslant v] \ = \ [Y \leqslant u] \ \cup \ [u < Y \leqslant v]$

et
$$\begin{aligned}[X \leqslant Y] \cap [Y \leqslant v] &= [X \leqslant Y] \cap \Big([Y \leqslant u] \ \cup \ [u < Y \leqslant v]\Big) \\ &= [X \leqslant Y] \cap [Y \leqslant u] \ \cup \ [X \leqslant Y] \cap [u < Y \leqslant v]\end{aligned}$$

enfin
$$\begin{aligned}\mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant v]\big) &= \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant u] \ \cup \ [X \leqslant Y] \cap [u < Y \leqslant v]\big) \\ &= \mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant u]\big) + \mathbb{P}\big([X \leqslant Y] \cap [u < Y \leqslant v]\big)\end{aligned}$$

La dernière égalité est obtenue par incompatibilité des deux événements considérés. En effet :

$$[Y \leqslant u] \cap [u < Y \leqslant v] = \varnothing$$

et ainsi : $[X \leqslant Y] \cap [Y \leqslant u] \ \cap \ [X \leqslant Y] \cap [u < Y \leqslant v] = \varnothing$.

$$\boxed{\text{On a bien : } H(v) - H(u) = \mathbb{P}\big(\ [X \leqslant Y] \cap [u < Y \leqslant v]\ \big).}$$

- Remarquons alors :

$$[X \leqslant Y] \cap [u < Y \leqslant v] \ \subset \ [X \leqslant v] \cap [u < Y \leqslant v]$$

Démontrons cette inclusion.

Soit $\omega \in \Omega$. Supposons $\omega \in [X \leqslant Y] \cap [u < Y \leqslant v]$.

On en déduit	$\omega \in [X \leqslant Y]$ et $\omega \in [u < Y \leqslant v]$.
Autrement dit	$X(\omega) \leqslant Y(\omega)$ et $u < Y(\omega) \leqslant v$.
En particulier	$X(\omega) \leqslant Y(\omega) \leqslant v$, ce qui s'écrit : $\omega \in [X \leqslant v]$.
Finalement	$\omega \in [X \leqslant v] \cap [u < Y \leqslant v]$.

$$\boxed{[X \leqslant Y] \cap [u < Y \leqslant v] \ \subset \ [X \leqslant v] \cap [u < Y \leqslant v]}$$

En particulier, par croissance de l'application $\mathbb{P}$, on obtient :

$$\mathbb{P}\big([X \leqslant Y] \cap [u < Y \leqslant v]\big) \ \leqslant \ \mathbb{P}\big([X \leqslant v] \cap [u < Y \leqslant v]\big)$$

- On a alors :

$$\begin{aligned} H(v) - H(u) &= \mathbb{P}\big([X \leqslant Y] \cap [u < Y \leqslant v]\big) \\ &\leqslant \mathbb{P}\big([X \leqslant v] \cap [u < Y \leqslant v]\big) \\ &= \mathbb{P}\big([X \leqslant v]\big) \times \mathbb{P}\big([u < Y \leqslant v]\big) && \textit{(car les v.a.r. X et Y sont indépendantes)} \\ &= F_X(v) \times \big(F_Y(v) - F_Y(u)\big) \end{aligned}$$

$$\boxed{\text{En divisant par } v - u > 0, \text{ on obtient bien : } \dfrac{H(v) - H(u)}{v - u} \leqslant F_X(v)\ \dfrac{F_Y(v) - F_Y(u)}{v - u}.}$$

- On raisonne de même pour l'inégalité de gauche. On établit initialement l'égalité :

$$[X \leqslant u] \cap [u < Y \leqslant v] \ \subset \ [X \leqslant Y] \cap [u < Y \leqslant v]$$

Soit $\omega \in \Omega$. Supposons $\omega \in [X \leqslant u] \cap [u < Y \leqslant v]$.

On en déduit	$\omega \in [X \leqslant u]$ et $[u < Y \leqslant v]$.
Autrement dit	$X(\omega) \leqslant u$ et $u < Y(\omega) \leqslant v$.
En particulier	$X(\omega) \leqslant u < Y(\omega)$, ce qui démontre : $\omega \in [X \leqslant Y]$.
Finalement	$\omega \in [X \leqslant Y] \cap [u < Y \leqslant v]$.

$$\boxed{\begin{array}{c}\text{On conclut alors, par un raisonnement similaire au précédent :} \\ F_X(u)\ \dfrac{F_Y(v) - F_Y(u)}{v - u} \ \leqslant \ \dfrac{H(v) - H(u)}{v - u}.\end{array}}$$

Commentaire

- Cette question peut sembler difficile car elle demande de nouveau une prise d'initiative importante. En particulier, il peut paraître difficile de penser à établir les inclusions entre événements qui permettent d'obtenir le résultat final. Il est conseillé d'opérer par rétro-ingénierie : on part du résultat final pour essayer d'en déduire le résultat intermédiaire qui permettra de conclure. Pour ce faire, on commence généralement par opérer par équivalence afin de pouvoir écrire le résultat final sous une forme plus simple. Par exemple, ici, on pouvait procéder comme suit :

$$\begin{array}{cll} & F_X(u)\ \dfrac{F_Y(v)-F_Y(u)}{v-u} \ \leqslant\ \dfrac{H(v)-H(u)}{v-u} & \\ \Leftrightarrow & F_X(u)\ \big(F_Y(v)-F_Y(u)\big) \ \leqslant\ H(v)-H(u) & \textit{(car } v-u>0\textit{)} \\ \Leftrightarrow & \mathbb{P}\big([X\leqslant u]\big)\ \mathbb{P}\big([u<Y\leqslant v]\big) \ \leqslant\ H(v)-H(u) & \textit{(par définition de } F_X \textit{ et propriété du cours)} \\ \Leftrightarrow & \mathbb{P}\big([X\leqslant u]\big)\ \mathbb{P}\big([u<Y\leqslant v]\big) \ \leqslant\ \mathbb{P}\big([X\leqslant Y]\cap[u<X\leqslant v]\big) & \\ \Leftrightarrow & \mathbb{P}\big([X\leqslant u]\cap[u<Y\leqslant v]\big) \ \leqslant\ \mathbb{P}\big([X\leqslant Y]\cap[u<X\leqslant v]\big) & \textit{(car } X \textit{ et } Y \textit{ sont indépendantes)} \end{array}$$

  Une inégalité entre probabilité est généralement obtenue par une inclusion entre événements. Il doit donc être naturel de penser à établir une telle inclusion. Notons enfin que l'on perd ici l'équivalence : l'inclusion entre événements **suffit** à démontrer l'inégalité entre probabilités.

- Au passage, soulignons de nouveau l'importance du triptyque :

intersection / indépendance / produit

  Lorsque la propriété établit une égalité qui comporte un produit de probabilités, il est naturel de penser que ce produit est obtenu comme probabilité d'une intersection d'événements.

- La difficulté d'un sujet se mesure en grande partie à la manière dont chaque question est découpée en sous-question. Mais il y a de sous-questions, plus le candidat doit prendre des initiatives. Ainsi, un sujet de type `TOP3` proposera un découpage en sous-questions bien moins détaillé qu'un sujet `TOP5`. Le même thème amène à un traitement différent lorsqu'il est abordé dans un sujet du `TOP3` ou du `TOP5`. En guise d'illustration, on peut noter que la propriété qu'il s'agit de démontrer dans cette **Partie I**, à savoir :

$$\mathbb{P}\big([X\leqslant Y]\big) \ = \ \int_0^{+\infty} F_X(t)\, f_Y(t)\ dt$$

  (sous les hypothèses de l'énoncé)

  est aussi utilisée (elle est admise) dans l'exercice 1 de l'énoncé `EML 2019`. □

b) En déduire que pour tout $x\in\mathbb{R}_+$, H est dérivable en x et : $H'(x) = F_X(x)\, f_Y(x)$.

Démonstration.

- Les fonctions f_X et f_Y sont continues sur l'intervalle $[0,+\infty[$.
  On en déduit que les fonctions F_X et F_Y sont de classe $\mathcal{C}^1$ sur cet intervalle.

En particulier, F_X et F_Y sont dérivables en tout point de $\mathbb{R}_+$.

  On en déduit que pour tout $x_0\in\mathbb{R}_+$: $\displaystyle\lim_{x\to x_0} \frac{F_Y(x)-F_Y(x_0)}{x-x_0} = F_Y'(x_0) = f_Y(x_0)$.

- Soit $x_0 \in \mathbb{R}_+$. Rappelons tout d'abord, que d'après la question précédente, pour tout $x > x_0$:

$$F_X(x_0)\ \frac{F_Y(x) - F_Y(x_0)}{x - x_0} \ \leqslant\ \frac{H(x) - H(x_0)}{x - x_0} \ \leqslant\ F_X(x)\ \frac{F_Y(x) - F_Y(x_0)}{x - x_0}$$

(résultat de la question précédente avec $v = x$ et $u = x_0$)

On a :

× $\lim\limits_{\substack{x \to x_0 \\ x > x_0}} F_X(x_0) = F_X(x_0)$,

et $\lim\limits_{\substack{x \to x_0 \\ x > x_0}} \dfrac{F_Y(x) - F_Y(x_0)}{x - x_0} = F'_Y(x_0) = f_Y(x_0)$ car F_Y est dérivable en x_0.

× $\lim\limits_{\substack{x \to x_0 \\ x > x_0}} F_X(x) = F_X(x_0)$ car F_X est continue (à droite) en x_0,

et $\lim\limits_{\substack{x \to x_0 \\ x > x_0}} \dfrac{F_Y(x) - F_Y(x_0)}{x - x_0} = F'_Y(x_0) = f_Y(x_0)$ car F_Y est dérivable en x_0.

On en déduit, par théorème d'encadrement que la fonction H admet une limite à droite à droite en x_0, donnée par :

$$\lim_{\substack{x \to x_0 \\ x > x_0}} \frac{H(x) - H(x_0)}{x - x_0} \ =\ H'_d(x_0) \ =\ F_X(x_0)\ f_Y(x_0)$$

En utilisant de nouveau le résultat de la question précédente, on obtient, pour tout $x < x_0$:

$$F_X(x)\ \frac{F_Y(x_0) - F_Y(x)}{x_0 - x} \ \leqslant\ \frac{H(x_0) - H(x)}{x_0 - x} \ \leqslant\ F_X(x_0)\ \frac{F_Y(x_0) - F_Y(x)}{x_0 - x}$$

(résultat de la question précédente avec $v = x_0$ et $u = x$)

Ce qui s'écrit (en multipliant chaque quotient par $\frac{-1}{-1}$) :

$$F_X(x)\ \frac{F_Y(x) - F_Y(x_0)}{x - x_0} \ \leqslant\ \frac{H(x) - H(x_0)}{x - x_0} \ \leqslant\ F_X(x_0)\ \frac{F_Y(x) - F_Y(0)}{x - x_0}$$

On en déduit alors, en utilisant de nouveau par le théorème d'encadrement, que la fonction H admet une limite à gauche en x_0, donnée par :

$$\lim_{\substack{x \to x_0 \\ x < x_0}} \frac{H(x) - H(x_0)}{x - x_0} \ =\ H'_g(x_0) \ =\ F_X(x_0)\ f_Y(x_0)$$

- Finalement, la fonction H est dérivable à droite et à gauche en x_0. De plus :

$$H'_g(x_0) \ =\ F_X(x_0)\ f_Y(x_0) \ =\ H'_d(x_0)$$

$$\boxed{\text{Ainsi, pour tout } x_0 \in \mathbb{R}_+, \text{ la fonction } H \text{ est dérivable en } x_0 \text{ et } H'(x_0) = F_X(x_0)\ f_Y(x_0).}$$ □

c) En conclure que pour tout x réel positif : $H(x) = \displaystyle\int_0^x F_X(t)\, f_Y(t)\ dt$.

Démonstration.
Dans les questions précédentes, on a établi que la fonction H :

× est dérivable sur $\mathbb{R}_+$,

× admet pour dérivée sur $\mathbb{R}_+$ la fonction $h : t \mapsto F_X(t)\, f_Y(t)$,

× vérifie : $H(0) = 0$.

On en déduit que la fonction H est la primitive sur $\mathbb{R}_+$ et qui s'annule en 0 de la fonction h.

$$\boxed{\begin{array}{c}\text{En conclusion, la fonction } H \text{ est telle que :}\\ \forall x \in \mathbb{R}_+,\ H(x) = \displaystyle\int_0^x h(t)\ dt = \int_0^x F_X(t)\, f_Y(t)\ dt.\end{array}}$$

Commentaire

On est confronté ici à une question bilan qui consiste simplement à rappeler puis utiliser certains résultats précédents. Ces résultats étant fournis par l'énoncé, cette question peut être traitée même si les questions précédentes ne l'ont pas été. Il faut s'habituer à repérer ces questions qui permettent de prendre facilement des points.

□

3. Démontrer : $\mathbb{P}(\,[X \leqslant Y]\,) = \displaystyle\int_0^{+\infty} F_X(t)\, f_Y(t)\ dt$.

Démonstration.
Il suffit de remarquer :

$$\begin{array}{rcll}\mathbb{P}\big([X \leqslant Y]\big) & = & \displaystyle\lim_{x \to +\infty} H(x) & \textit{(d'après la question \textbf{1.})}\\[2mm] & = & \displaystyle\lim_{x \to +\infty} \int_0^x F_X(t)\, f_Y(t)\ dt & \textit{(d'après la question \textbf{2.c)})}\\[2mm] & = & \displaystyle\int_0^{+\infty} F_X(t)\, f_Y(t)\ dt & \end{array}$$

Rappelons que l'on a démontré, en question ***1.***, que la fonction $x \mapsto \displaystyle\int_0^x F_X(t)\, f_Y(t)\ dt$ admet une limite finie en $+\infty$ ($= \mathbb{P}\big([X \leqslant Y]\big)$). Cela signifie, par définition, que l'intégrale impropre $\displaystyle\int_0^{+\infty} F_X(t)\, f_Y(t)\ dt$ est convergente. Cela justifie la dernière égalité.

$$\boxed{\mathbb{P}(\,[X \leqslant Y]\,) = \int_0^{+\infty} F_X(t)\, f_Y(t)\ dt}$$

Commentaire

- Il s'agit là encore d'une question bilan qui ne présente pas de difficulté particulière. Cela démontre au passage qu'il n'y a pas forcément de progression croissante de la difficulté des questions dans les énoncés des épreuves de concours. En conséquence, même si on ne parvient pas à traiter plusieurs questions d'affilée, il ne faut pas pour autant passer toute la **Partie I**. Il faut au contraire s'atteler à essayer de traiter les questions qui suivent, ce qui permettra à terme de tomber sur une question dont la résolution est plus simple.

Commentaire

- Dans l'épreuve `EML 2019`, on admet l'écriture :
$$\mathbb{P}\big([X \leqslant Y]\big) = \int_0^{+\infty} F_X(t)\, f_Y(t)\, dt$$
Il est par contre demandé de justifier la convergence de cette intégrale à l'aide d'un théorème de comparaison. Rappelons cette démonstration :
  × $\forall t \in [0, +\infty[,\ 0 \leqslant F_X(t)\, f_Y(t) \leqslant f_Y(t)$
  En effet, pour tout $t \in [0, +\infty[$, $0 \leqslant F_X(t) \leqslant 1$ et l'inégalité souhaitée est alors obtenue par multiplication par $f_Y(t) \geqslant 0$.
  × l'intégrale $\displaystyle\int_0^{+\infty} f_Y(t)\, dt$ est convergente (et vaut 1) en tant que moment d'ordre 0 de la v.a.r. Y. En effet, comme Y est à valeurs positives, f_Y est nulle en dehors de $[0, +\infty[$ et :
$$\mathbb{E}(Y^0) = \int_{-\infty}^{+\infty} f_Y(t)\, dt = \int_0^{+\infty} f_Y(t)\, dt$$
Ainsi, par critère de comparaison d'intégrales généralisées de fonctions continues positives, l'intégrale impropre $\displaystyle\int_0^{+\infty} F_X(t)\, f_Y(t)\, dt$ est convergente.
- On pouvait aussi opérer à l'aide d'un équivalent. En effet, comme : $\lim\limits_{t \to +\infty} F_X(t) = 1$, on a :
$$F_X(t)\, f_Y(t) \underset{t \to +\infty}{\sim} f_Y(t)$$

□

4. En utilisant la fonction $K : x \mapsto \mathbb{P}([X < Y] \cap [Y \leqslant x])$, on montrerait de même et nous l'admettrons :
$$\mathbb{P}([X < Y]) = \int_0^{+\infty} F_X(t)\, f_Y(t)\, dt = \mathbb{P}([X \leqslant Y])$$
Que peut-on en déduire pour $\mathbb{P}([X = Y])$?

Démonstration.

- Remarquons tout d'abord :
$$[X \leqslant Y] = [X < Y] \cup [X = Y]$$
- On en déduit :
$$\begin{aligned} \mathbb{P}\big([X \leqslant Y]\big) &= \mathbb{P}\big([X < Y] \cup [X = Y]\big) \\ &= \mathbb{P}\big([X < Y]\big) + \mathbb{P}\big([X = Y]\big) \end{aligned}$$
(car les événements $[X < Y]$ et $[X = Y]$ sont incompatibles)

$$\boxed{\text{Ainsi : } \mathbb{P}\big([X = Y]\big) = \mathbb{P}\big([X \leqslant Y]\big) - \mathbb{P}\big([X < Y]\big) = 0.}$$

□

5. *Application aux lois exponentielles*

On suppose que U et V sont deux variables aléatoires indépendantes suivant des lois exponentielles de paramètres respectifs λ et μ, réels strictement positifs.

Soit θ un réel positif ou nul.

a) Déterminer la fonction de répartition de la variable aléatoire $X = U - \theta$.

Démonstration.

- Notons $h_\theta : x \mapsto x - \theta$ de sorte que $X = h_\theta(U)$.

  Comme $U \hookrightarrow \mathcal{E}(\lambda)$, on **considère** $X(\Omega) = [0, +\infty[$. On en déduit :

$$\begin{array}{rcll} X(\Omega) & = & \big(h_\theta(U)\big)(\Omega) & \\ & = & h_\theta\big(U(\Omega)\big) & \\ & = & h_\theta\big([0, +\infty[\big) & \\ & = & [h_\theta(0), \lim\limits_{x \to +\infty} h_\theta(x)[& \textit{(car la fonction } h_\theta \textit{ est continue et strictement croissante sur } [0, +\infty[\textit{)} \\ & = & [-\theta, +\infty[& \end{array}$$

$$\boxed{\text{Et ainsi : } X(\Omega) = [-\theta, +\infty[.}$$

- Soit $x \in \mathbb{R}$. Deux cas se présentent :

  × si $x < -\theta$, alors $[X \leqslant x] = \varnothing$ car $X(\Omega) = [-\theta, +\infty[$. Donc :

$$F_X(x) = \mathbb{P}\big(\,[X \leqslant x]\,\big) = \mathbb{P}(\varnothing) = 0$$

  × si $x \geqslant -\theta$, alors :

$$\begin{array}{rcll} F_X(x) & = & \mathbb{P}\big(\,[X \leqslant x]\,\big) & \\ & = & \mathbb{P}\big(\,[U - \theta \leqslant x]\,\big) & \\ & = & \mathbb{P}\big(\,[U \leqslant x + \theta]\,\big) & \\ & = & 1 - \exp\big(-\lambda\,(x + \theta)\big) & \textit{(car } X \hookrightarrow \mathcal{E}(\lambda) \textit{ et } x + \theta \geqslant 0\textit{)} \end{array}$$

$$\boxed{\text{On obtient finalement : } F_X : x \mapsto \left\{ \begin{array}{cl} 0 & \text{si } x < -\theta \\ 1 - \mathrm{e}^{-\lambda\,\theta}\, \mathrm{e}^{-\lambda\, x} & \text{si } x \geqslant -\theta \end{array} \right. .}$$

Commentaire

- Cette question consiste à déterminer la loi de Y, transformée de la v.a.r. X. Ce type de question est extrêmement fréquent dans les sujets traitant de v.a.r. à densité. La résolution de ce type de question ne présente aucune difficulté majeure. Il s'agit simplement de se référer à la rédaction usuelle.
- En particulier, il faut savoir déterminer la loi d'une transformée affine, du carré et de la partie entière d'une v.a.r. à densité X. Cela fait partie du bagage culturel mathématique nécessaire avant d'affronter les écrits de concours.
- Il faut ajouter à ce bagage la détermination de la loi du minimum et du maximum de v.a.r. à densité indépendantes. Il suffit une nouvelle fois de mettre en place la rédaction usuelle associée à ce type de questions.

Commentaire

- Profitons-en pour faire un point sur la notation $X(\Omega)$.
  Rappelons qu'une v.a.r. X est une application $X : \Omega \to \mathbb{R}$.
  Comme la notation le suggère, $X(\Omega)$ est l'image de Ω par l'application X.
  Ainsi, $X(\Omega)$ n'est rien d'autre que l'ensemble des valeurs prises par la v.a.r. X :
  $$\begin{aligned} X(\Omega) &= \{X(\omega) \mid \omega \in \Omega\} \\ &= \{x \in \mathbb{R} \mid \exists \omega \in \Omega,\ X(\omega) = x\} \end{aligned}$$
  Il faut bien noter que dans cette définition aucune application probabilité $\mathbb{P}$ n'apparaît.
- Il est toujours correct d'écrire : $X(\Omega) \ \subseteq\ ]-\infty, +\infty[$.
  En effet, cette propriété signifie que toute v.a.r. X est à valeurs dans $\mathbb{R}$, ce qui est toujours le cas par définition de la notion de variable aléatoire réelle.
- Dans le cas des v.a.r. discrètes, il est d'usage relativement courant de confondre :
  - × l'ensemble des valeurs possibles de la v.a.r. X (*i.e.* l'ensemble $X(\Omega)$),
  - × l'ensemble $\{x \in \mathbb{R} \mid \mathbb{P}\big([X = x]\big) \neq 0\}$, ensemble des valeurs que X prend avec probabilité non nulle. Dans le cas qui nous intéresse ici, à savoir X est une v.a.r. discrète, cet ensemble est appelé support de X et est noté $\mathrm{Supp}(X)$.
- Dans le cas des v.a.r. à densité, la détermination de l'ensemble image est plus technique. Dans certains sujets, l'ensemble image des v.a.r. étudiées sera précisé (« On considère une v.a.r. à valeurs strictement positives »). Si ce n'est pas le cas :
  - × si X suit une loi usuelle, on peut se référer à l'ensemble image donné en cours. Par exemple, si $X \hookrightarrow \mathcal{U}([0, 1])$, on se permet d'écrire :

    « Comme $X \hookrightarrow \mathcal{U}([0, 1])$, on **considère** : $X(\Omega) = [0, 1]$. »

  - × si X ne suit pas une loi usuelle, on étudie l'ensemble : $I = \{x \in \mathbb{R} \mid f_X(x) > 0\}$.
    On se permet alors d'écrire :

    « Dans la suite, on **considère** : $X(\Omega) = I$. »

  En **décrétant** la valeur de $X(\Omega)$, on ne commet pas une erreur mais on décide d'ajouter une hypothèse qui ne fait pas partie de l'énoncé. Cette audace permet de travailler avec un ensemble image connu, ce qui permet de structurer certaines démonstrations (l'ensemble image étant connu, on se rappelle que la fonction de répartition, par exemple, s'obtient par une disjonction de cas).

□

b) En déduire que pour tout $\theta \geqslant 0$:
$$\mathbb{P}(\,[U - \theta \leqslant V]\,) \ = \ 1 - \frac{\mu}{\lambda + \mu}\ \mathrm{e}^{-\lambda\,\theta}$$

Démonstration.

- Remarquons tout d'abord que la v.a.r. $X = U - \theta$ est une v.a.r. à densité en tant que transformée affine d'une v.a.r. à densité.
- Ainsi, on a :
  - × les v.a.r. X et V sont à densité.
  - × les v.a.r. $X = U - \theta$ et V sont indépendantes d'après le lemme des coalitions et car U et V le sont.

× comme $V \hookrightarrow \mathcal{E}(\mu)$, on peut considérer $V(\Omega) = [0, +\infty[$.
Autrement dit, on considère que la v.a.r. V est à valeurs positives.

× une densité f_V de la v.a.r. V est donnée par :

$$f_V : t \mapsto \begin{cases} 0 & \text{si } t < 0 \\ \mu \, \mathrm{e}^{-\mu \, t} & \text{si } t \geqslant 0 \end{cases}$$

Et ainsi : $f_V|_{[0,+\infty[} : t \mapsto \mu \, \mathrm{e}^{-\mu \, t}$ est une fonction continue sur $[0, +\infty[$.

On est dans le cadre d'application du résultat démontré en question ***4***.

Commentaire

- Comme précisé dans la question précédente, on se permet de considérer $V(\Omega) = [0, +\infty[$. En réalité, pour toute v.a.r. qui suit une loi exponentielle, cette propriété n'est vérifiée que presque sûrement (avec probabilité 1). Autrement dit, on a toujours, sans hypothèse supplémentaire : $\mathbb{P}\big([V \geqslant 0]\big) = 1$. Il est à noter que c'est l'énoncé qui nous amène à considérer $V(\Omega) = [0, +\infty[$. En effet, cette hypothèse est nécessaire pour se placer dans le cadre d'application du résultat démontré en question ***4***. Il aurait donc été préférable que l'énoncé précise que V est une v.a.r. à valeurs positives, en début de question ***5***.
- Comme signalé au-dessus, il est primordial de savoir déterminer la transformée affine d'une v.a.r. à densité. Ici, on ne demande pas explicitement d'obtenir une densité de la v.a.r. X. On utilise ici le résultat du cours qui affime que la transformée affine d'une v.a.r. à densité est une v.a.r. à densité. On peut aussi démontrer que X est une v.a.r. à densité en établissant que F_X est :
  × continue sur $\mathbb{R}$,
  × de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf (éventuellement) en un nombre fini de points.

- On a alors :

$$\begin{aligned}
\mathbb{P}\big([U - \theta \leqslant V]\big) &= \mathbb{P}\big([X \leqslant V]\big) \\
&= \int_0^{+\infty} F_X(t) \, f_V(t) \, dt \\
&= \int_0^{+\infty} \big(1 - \mathrm{e}^{-\lambda \theta} \, \mathrm{e}^{-\lambda t}\big) \, \mu \, \mathrm{e}^{-\mu t} \, dt \\
&= \int_0^{+\infty} \mu \, \mathrm{e}^{-\mu t} \, dt - \int_0^{+\infty} \mathrm{e}^{-\lambda \theta} \, \mathrm{e}^{-\lambda t} \, \mu \, \mathrm{e}^{-\mu t} \, dt && \textit{(par linéarité de l'intégration, les intégrales en présence étant convergentes)} \\
&= \mathbb{E}\big(V^0\big) - \mathrm{e}^{-\lambda \theta} \, \mu \int_0^{+\infty} \mathrm{e}^{-\lambda t} \, \mathrm{e}^{-\mu t} \, dt \\
&= 1 - \mathrm{e}^{-\lambda \theta} \, \mu \, \frac{1}{\lambda + \mu} \int_0^{+\infty} (\lambda + \mu) \, \mathrm{e}^{-(\lambda+\mu) \, t} \, dt \\
&= 1 - \frac{\mu}{\lambda + \mu} \, \mathrm{e}^{-\lambda \theta} \, \mathbb{E}\big(W^0\big) && \textit{(où } W \textit{ est une v.a.r. de loi } \mathcal{E}(\lambda + \mu)\textit{)}
\end{aligned}$$

$$\boxed{\text{Finalement, on obtient bien : } \mathbb{P}\big([U - \theta \leqslant V]\big) = 1 - \frac{\mu}{\lambda + \mu} \, \mathrm{e}^{-\lambda \theta}.}$$

□

Inégalité de Boole

6. On considère $(B_k)_{k\in\mathbb{N}^*}$ une famille d'événements d'un espace probabilisé.

a) Montrer par récurrence sur $n \in \mathbb{N}^*$: $\mathbb{P}\left(\bigcup\limits_{k=1}^{n} B_k\right) \leqslant \sum\limits_{k=1}^{n} \mathbb{P}(B_k)$.

Démonstration.
Démontrons par récurrence : $\forall n \in \mathbb{N}^*$, $\mathcal{P}(n)$, où $\mathcal{P}(n)$: $\mathbb{P}\left(\bigcup\limits_{k=1}^{n} B_k\right) \leqslant \sum\limits_{k=1}^{n} \mathbb{P}(B_k)$.

- **Initialisation** :
  - D'une part : $\mathbb{P}\left(\bigcup\limits_{k=1}^{1} B_k\right) = \mathbb{P}(B_1)$.
  - D'autre part : $\sum\limits_{k=1}^{1} \mathbb{P}(B_k) = \mathbb{P}(B_1)$

  On a bien : $\mathbb{P}(B_1) \leqslant \mathbb{P}(B_1)$. D'où $\mathcal{P}(1)$.
- **Hérédité** : soit $n \in \mathbb{N}^*$.
  Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$ $\left(i.e. : \mathbb{P}\left(\bigcup\limits_{k=1}^{n+1} B_k\right) \leqslant \sum\limits_{k=1}^{n+1} \mathbb{P}(B_k)\right)$.

$$\begin{array}{rcll} \mathbb{P}\left(\bigcup\limits_{k=1}^{n+1} B_k\right) & = & \mathbb{P}\left(\left(\bigcup\limits_{k=1}^{n} B_k\right) \cup B_{n+1}\right) & \\ & = & \mathbb{P}\left(\bigcup\limits_{k=1}^{n} B_k\right) + \mathbb{P}(B_{n+1}) - \mathbb{P}\left(\left(\bigcup\limits_{k=1}^{n} B_k\right) \cap B_{n+1}\right) & \textit{(d'après la formule du crible)} \\ & = & \sum\limits_{k=1}^{n+1} \mathbb{P}(B_k) + \mathbb{P}(B_{n+1}) - \mathbb{P}\left(\left(\bigcup\limits_{k=1}^{n} B_k\right) \cap B_{n+1}\right) & \textit{(par hypothèse de récurrence)} \\ & \leqslant & \sum\limits_{k=1}^{n+1} \mathbb{P}(B_k) + \mathbb{P}(B_{n+1}) & \end{array}$$

  D'où $\mathcal{P}(n+1)$.

$$\boxed{\text{Par principe de récurrence : } \forall n \in \mathbb{N}^*,\ \mathbb{P}\left(\bigcup\limits_{k=1}^{n} B_k\right) \leqslant \sum\limits_{k=1}^{n} \mathbb{P}(B_k).}$$

□

b) On suppose que la série $\sum\limits_{k\geqslant 1} \mathbb{P}(B_k)$ converge. Démontrer :

$$\mathbb{P}\left(\bigcup\limits_{k=1}^{+\infty} B_k\right) \leqslant \sum\limits_{k=1}^{+\infty} \mathbb{P}(B_k)$$

Démonstration.

- Tout d'abord, par le théorème de la limite monotone : $\mathbb{P}\left(\bigcup\limits_{k=1}^{+\infty} B_k\right) = \lim\limits_{N\to+\infty} \mathbb{P}\left(\bigcup\limits_{k=1}^{N} B_k\right)$.
- Comme la série $\sum\limits_{k\geqslant 1} \mathbb{P}(B_k)$ est supposée convergente, on obtient, par passage à la limite dans l'inégalité de la question précédente :

$$\lim\limits_{N\to+\infty} \mathbb{P}\left(\bigcup\limits_{k=1}^{N} B_k\right) \leqslant \lim\limits_{N\to+\infty} \sum\limits_{k=1}^{N} \mathbb{P}(B_k) = \sum\limits_{k=1}^{+\infty} \mathbb{P}(B_k)$$

$$\boxed{\mathbb{P}\left(\bigcup\limits_{k=1}^{+\infty} B_k\right) \leqslant \sum\limits_{k=1}^{+\infty} \mathbb{P}(B_k)}$$

□

Partie II - Une compétition entre deux groupes

Dans toute la suite du sujet, on désigne par p un réel de l'intervalle $]0,1[$ et on pose $q = 1 - p$.
On modélise une compétition entre deux groupes d'individus A et B avec les règles suivantes.

- Le groupe A doit résoudre une suite de problèmes $(P_k)_{k\geqslant 1}$ dans l'ordre des indices. Au temps $t = 0$, le groupe commence la résolution du problème P_1, ce qui lui prend un temps représenté par la variable aléatoire X_1. Une fois P_1 résolu, le groupe aborde immédiatement le problème P_2, et on note X_2 le temps consacré à la résolution de P_2 par le groupe A, et ainsi de suite.
  Pour tout $k \in \mathbb{N}^*$, on note X_k la variable aléatoire donnant le temps consacré à la résolution du problème P_k par le groupe A.
- De même, le groupe B doit résoudre dans l'ordre une suite de problèmes $(Q_k)_{k\geqslant 1}$; la résolution du premier problème Q_1 commence au temps $t = 0$ et on note, pour tout $k \in \mathbb{N}^*$, Y_k la variable aléatoire donnant le temps consacré par le groupe B à la résolution du problème Q_k.
- À ce jeu est associé un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ sur lequel sont définies les suites de variables aléatoires $(X_k)_{k\geqslant 1}$ et $(Y_k)_{k\geqslant 1}$, et on fait les hypothèses suivantes :
  - × pour tout $k \in \mathbb{N}^*$, X_k suit la loi exponentielle de paramètre p, notée $\mathcal{E}(p)$, et Y_k suit la loi exponentielle $\mathcal{E}(q)$;
  - × pour tout $k \in \mathbb{N}^*$, les variables aléatoires $X_1, \ldots, X_k, Y_1, \ldots, Y_k$ sont indépendantes.
- On établit alors la liste de tous les problèmes résolus *dans l'ordre où ils le sont par les deux groupes.* En cas de simultanéité temporelle de la résolution par les deux groupes d'un de leurs problèmes, on placera d'abord le problème résolu par A dans la liste puis celui résolu par B.
  Pour tout $n \in \mathbb{N}^*$, on note U_n la variable aléatoire de Bernoulli associée à l'événement « le $n^{\text{ème}}$ problème placé dans la liste est un problème résolu par le groupe A ».
  Par exemple, si la liste des cinq premiers problèmes résolus est $(P_1, P_2, Q_1, P_3, Q_2)$, alors $U_1 = 1$, $U_2 = 1$, $U_3 = 0$, $U_4 = 1$ et $U_5 = 0$.
- Pour tout $n \geqslant 0$, on note aussi S_n la variable aléatoire donnant le nombre de problèmes qui ont été résolus par A présents dans la liste des n premiers problèmes résolus. En particulier, S_0 vaut toujours 0.

7. ***a)*** Que représente la variable aléatoire $\sum\limits_{k=1}^{n} X_k$?

b) On suppose que $X_1 = 5$, $X_2 = 2$, $X_3 = 3$, $X_4 = 2$, $Y_1 = 2$, $Y_2 = 2$, $Y_3 = 4$ et $Y_4 = 2$.
Déterminer U_1, ..., U_7.
Peut-on aussi en déduire la valeur de U_8 ?

c) Compléter le script **Python** suivant pour qu'il simule le jeu et, pour n, p donnés, affiche la liste des valeurs U_1, U_2, ..., U_n :

```
p = float(input('p = '))
n = int(input('n = '))
q = 1 - p
U = np.zeros(n)
sommeX = rd.exponential(1/p)
sommeY = rd.exponential(1/q)
mini = min(sommeX, sommeY)
for k in range(n):
    if sommeX == ...:
        U[k] = ...
        sommeX = sommeX + rd.exponential(1/p)
    else:
        sommeY = ...
    mini = min(sommeX, sommeY)
...
```

Démonstration.
On propose le script suivant :

```
p = float(input('p = '))
n = int(input('n = '))
q = 1 - p
U = np.zeros(n)
sommeX = rd.exponential(1/p)
sommeY = rd.exponential(1/q)
mini = min(sommeX, sommeY)
for k in range(n):
    if sommeX == mini:
        U[k] = 1
        sommeX = sommeX + rd.exponential(1/p)
    else:
        sommeY = sommeY + rd.exponential(1/q)
    mini = min(sommeX, sommeY)
print(U)
```

□

d) Quelle(s) instruction(s) faut-il ajouter pour afficher la valeur de S_n ?

8. *Loi de U_n*
Dans cette question, on démontre par récurrence sur $n \geqslant 1$: $\mathbb{P}([U_n = 1]) = p$.

a) Démontrer : $\mathbb{P}([U_1 = 1]) = \mathbb{P}([X_1 \leqslant Y_1]) = p$.

b) (i) Démontrer, pour tout réel $x < 0$: $\mathbb{P}_{[U_1=1]}([Y_1 - X_1 \leqslant x]) = 0$.

(ii) Soit x un réel positif ou nul.
Établir : $\mathbb{P}_{[U_1=1]}(\,[Y_1 - X_1 \leqslant x]\,) = \dfrac{1}{p}\ \mathbb{P}(\,[X_1 \leqslant Y_1 \leqslant X_1 + x]\,)$,
puis calculer $\mathbb{P}_{[U_1=1]}(\,[Y_1 - X_1 \leqslant x]\,)$.

c) On peut interpréter ce résultat en disant que la *loi conditionnelle de* $Y_1 - X_1$ *sachant* $[U_1 = 1]$ est une loi exponentielle. Quelle est son paramètre ?
Par analogie, quelle est la loi conditionnelle de $X_1 - Y_1$ sachant $[U_1 = 0]$? (on n'attend pas une démonstration précise mais un argument de bon sens pour justifier le résultat proposé).

d) On suppose que $n \in \mathbb{N}^*$ et $\mathbb{P}(\,[U_n = 1]\,) = p$.
Déduire de cette hypothèse et de la question précédente :

$$\mathbb{P}_{[U_1=1]}(\,[U_{n+1} = 1]\,) = p \quad \text{et} \quad \mathbb{P}_{[U_1=0]}(\,[U_{n+1} = 1]\,) = \ p$$

e) Conclure.

9. On montrerait aussi par récurrence, et nous l'admettrons, que pour tout $n \in \mathbb{N}^*$, les variables aléatoires $U_1, \ldots, U_n$ sont mutuellement indépendantes.
En déduire la loi de S_n.

Démonstration.
On remarque que $S_n = \sum\limits_{k=1}^{n} U_k$. De plus :

- les variables aléatoires U_k sont indépendantes
- pour tout $k \in [\![1, n]\!]$, $U_k \hookrightarrow \mathcal{B}\,(p)$

$$\boxed{\text{Ainsi, par théorème de stabilité des lois binomiales : } S_n \hookrightarrow \mathcal{B}\,(n, p).}$$

□

Soit $r \in \mathbb{N}$, on s'intéresse, dans les questions qui suivent, à la probabilité a_r de l'événement :

A_r : « il existe un $n \geqslant r$ tel que, lorsque n problèmes en tout ont été résolus, le groupe A en a résolu r de plus que le groupe B »

10. ***a)*** Justifier : $a_0 = 1$.

b) Démontrer, pour tout $r \geqslant 1$:

$$\mathbb{P}_{[U_1=1]}(A_r) = \mathbb{P}(A_{r-1}) \quad \text{et} \quad \mathbb{P}_{[U_1=0]}(A_r) = \mathbb{P}(A_{r+1})$$

c) En déduire, pour tout $r \geqslant 1$: $a_{r+1} = \dfrac{1}{q}\, a_r - \dfrac{p}{q}\, a_{r-1}$.

d) En remarquant que $1 - 4pq = (1-2p)^2$, donner une expression de a_r en fonction de p, q, r et de deux constantes que l'on introduira.

11. Le cas $p \geqslant \dfrac{1}{2}$.

Montrer que, dans les cas $p = \dfrac{1}{2}$ et $p > \dfrac{1}{2}$, la suite $(a_r)_{r\in\mathbb{N}}$ est constante et égale à 1.

12. Le cas $p < \dfrac{1}{2}$.

a) Soit k un entier naturel.

(i) Établir : $A_{2k} = \bigcup\limits_{i \geqslant k} [S_{2i} = i + k]$.

(ii) Montrer que pour tout $i \geqslant k$, on a : $\mathbb{P}(\,[S_{2i} = i+k]\,) = \dbinom{2i}{i+k} p^{i+k}\, q^{i-k}$.

(iii) Après avoir donné la valeur de la somme $\sum\limits_{j=0}^{2i} \dbinom{2i}{j}$, démontrer :

$$\forall i \geqslant k,\ \binom{2i}{i+k} \leqslant 4^i$$

(iv) En déduire l'inégalité :

$$\sum_{i=k}^{+\infty} \mathbb{P}(\,[S_{2i} = k+i]\,) \ \leqslant\ \left(\frac{p}{q}\right)^k \ \frac{(4pq)^k}{1-4pq}$$

b) Montrer en utilisant l'inégalité de Boole (voir question ***6.***) que si $p < \dfrac{1}{2}$, alors : $\lim\limits_{k\to+\infty} a_{2k} = 0$.

c) Conclure en utilisant la question ***10.d)***, que si $p < \dfrac{1}{2}$, alors :

$$\forall r \in \mathbb{N},\ a_r = \left(\frac{p}{q}\right)^r$$

On a ainsi établi dans les questions ***11.*** et ***12.*** :

$$\forall r \in \mathbb{N}, \quad a_r = \begin{cases} \left(\dfrac{p}{q}\right)^r & \text{si } p < \dfrac{1}{2} \\[2ex] 1 & \text{si } p \geqslant \dfrac{1}{2} \end{cases}$$

Ce résultat pourra être admis et utilisé dans la suite du sujet.

Partie III - La *blockchain* et la stratégie de la *double dépense*

On utilise, dans cette partie, les notations et résultats de la partie **II**.
Soit n un entier supérieur ou égal à 1.
La *blockchain* est formée d'une suite de blocs, chacun associé à plusieurs transactions. Elle contient l'historique de toutes les transactions effectuées depuis la création du *bitcoin*.
Avant d'être placé dans la *blockchain*, un nouveau bloc doit être validé. Cette validation nécessite la mise en oeuvre d'une grande puissance de calcul pour résoudre un problème dépendant fortement du contenu du bloc et des blocs qui le précèdent.
Les individus qui valident les blocs sont appelés mineurs.
Il est possible qu'à un instant donné, coexistent sur le réseau deux *blockchains*, valides et différentes. Dans ce cas, le réseau choisira celle qui comporte le plus de blocs et l'autre sera abandonnée.
Par prudence, lorsqu'un bloc est validé, il est recommandé d'attendre que $n-1$ blocs le suivant soient aussi validés pour considérer que les transactions incluses dans le bloc soient honnêtes.
Un groupe de mineurs mal intentionnés, noté A, peut essayer de dépenser deux fois les mêmes *bitcoins* en procédant ainsi :

- le groupe A demande la validation de l'achat d'un bien d'un montant de s *bitcoins* qu'il a en sa possession.
- lorsque le bloc K incluant cette transaction est proposé à la validation sur le réseau, A modifie ce bloc en K', qu'il ne diffuse pas, en remplaçant l'achat par une vente des s *bitcoins* en euros à son profit par exemple. Il se met alors à la validation de ce nouveau bloc et crée ainsi une deuxième instance de la *blockchain* qu'il continue à développer sans la diffuser.
- lorsque le groupe B, représentant l'ensemble des autres mineurs du réseau, a validé K ainsi que les $n-1$ blocs suivants, le vendeur du bien considère que la transaction est valide et fournit le bien.
- le groupe A attend alors d'avoir une *blockchain* plus longue que celle de B, qui est publique, pour la diffuser donc invalider la *blockchain* publique et l'achat du bien. Le crédit en *bitcoins* du vendeur du bien est alors annulé.

On reprend et on complète la modélisation de la partie précédente pour déterminer la probabilité que la stratégie de la *double dépense* réussisse et le choix de n pour que cette probabilité soit faible.
Une première phase du jeu, décrit dans la partie **II**, s'achève à l'instant aléatoire t où le problème Q_n est ajouté à la liste des problèmes résolus.
Le groupe de mineurs A est ensuite déclaré vainqueur s'il se trouve un instant $t' \geqslant t$ où le nombre de problèmes résolus par A dans la liste des problèmes résolus depuis le début du jeu, est strictement supérieur au nombre de ceux résolus par B dans cette même liste. On note G_n cet événement.
On détermine, dans cette partie, la probabilité de G_n en fonction de n et de p.

13. On s'intéresse tout d'abord à la loi de la variable aléatoire T_n égale au nombre de problèmes résolus par le groupe A lorsque l'on place Q_n dans la liste des problèmes résolus.

a) Démontrer, pour tout $k \in \mathbb{N}$: $[T_n = k] = [S_{n+k-1} = k] \cap [U_{n+k} = 0]$.

Démonstration.
Soit $k \in \mathbb{N}$. Soit $\omega \in \Omega$.

$$\begin{array}{cl} & \omega \in [S_{n+k-1} = k] \quad \cap \quad [U_{n+k} = 0] \\ \Leftrightarrow & \omega \in [S_{n+k-1} = k] \quad \texttt{ET} \quad \omega \in [U_{n+k} = 0] \\ \Leftrightarrow & S_{n+k-1}(\omega) = k \quad \texttt{ET} \quad U_{n+k}(\omega) = 0 \end{array}$$

On en déduit :

$$\omega \in [S_{n+k-1} = k] \cap [U_{n+k} = 0]$$

$$\Leftrightarrow \quad \begin{array}{c} \text{dans les } n+k-1 \text{ premiers problèmes résolus, } k \text{ l'ont été par le groupe } A \\ \text{et} \\ \text{le } (n+k)^{\text{ème}} \text{ problème a été résolu par le groupe } B \end{array}$$

$$\Leftrightarrow \quad \begin{array}{c} \text{sur les } n+k-1 \text{ premiers problèmes, } k \text{ ont été résolus par le groupe } A\text{, et} \\ n-1 \text{ l'ont été par le groupe } B \\ \text{et} \\ \text{le dernier problème (le } (n+k)^{\text{ème}}\text{) est résolu par le groupe } B \text{ (c'est donc le} \\ n^{\text{ème}} \text{ problème résolu par } B\text{, c'est-à-dire } Q_n\text{)} \end{array}$$

$$\Leftrightarrow \quad \begin{array}{c} \text{sur les } n+k-1 \text{ premiers problèmes, } k \text{ ont été résolus par le groupe } A \\ \text{et} \\ Q_n \text{ est le } (n+k)^{\text{ème}} \text{ problème résolu} \end{array}$$

$$\Leftrightarrow \quad \begin{array}{c} \text{lorsque } Q_n \text{ est résolu par le groupe } B\text{, } k \text{ problèmes ont été résolus par le} \\ \text{groupe } A \end{array}$$

$$\Leftrightarrow \quad T_n(\omega) = k$$

$$\Leftrightarrow \quad \omega \in [T_n = k]$$

$$\boxed{\text{On en déduit : } [T_n = k] = [S_{n+k-1} = k] \cap [U_{n+k} = 0].}$$

□

b) En déduire : $\mathbb{P}(\,[T_n = k]\,) = \dbinom{n+k-1}{k} p^k\, q^n$.

Démonstration.
Soit $k \in \mathbb{N}$.

- D'après la question précédente :

$$\mathbb{P}([T_n = k]) \;=\; \mathbb{P}([S_{n+k-1} = k] \cap [U_{n+k} = 0])$$

- Or, d'après la question ***9.*** : $S_{n+k-1} = \sum\limits_{i=1}^{n+k-1} U_i$.
  Ainsi, par lemme des coalitions, les v.a.r. S_{n+k-1} et U_{n+k} sont indépendantes.
- On en déduit :

$$\begin{aligned} \mathbb{P}([T_n = k]) &= \mathbb{P}([S_{n+k-1} = k]) \times \mathbb{P}([U_{n+k} = 0]) \\ &= \binom{n+k-1}{k} p^k\, q^{(n+\cancel{k}-1)-\cancel{k}} \times q \qquad \textit{(d'après les questions } \textbf{\textit{8.}} \textit{ et } \textbf{\textit{9.}}\textit{)} \\ &= \binom{n+k-1}{k} p^k\, q^n \end{aligned}$$

$$\boxed{\forall k \in \mathbb{N},\ \mathbb{P}(\,[T_n = k]\,) = \binom{n+k-1}{k} p^k\, q^n}$$

□

14. a) En utilisant la formule des probabilités totales, établir :

$$\mathbb{P}(G_n) \;=\; \mathbb{P}(\,[T_n \geqslant n+1]\,) + \sum_{k=0}^{n} \mathbb{P}(\,[T_n = k]\,)\, a_{n+1-k}$$

Démonstration.

- Le famille $\big(\,[T_n = k]\,\big)_{k\in\mathbb{N}}$ forme un système complet d'événements.
  Par formule des probabilités totales :

$$\begin{aligned}\mathbb{P}(G_n) &= \sum_{k=0}^{+\infty} \mathbb{P}([T_n = k] \cap G_n) \\ &= \sum_{k=0}^{+\infty} \mathbb{P}([T_n = k])\, \mathbb{P}_{[T_n=k]}(G_n) \quad \textit{(car : } \forall k \in \mathbb{N},\ \mathbb{P}([T_n = k]) = 0\textit{)}\end{aligned}$$

- Soit $k \in \mathbb{N}$.
  Si l'événement $[T_n = k]$ est réalisé, c'est que, lorsque Q_n est résolu par le groupe B, le groupe A a déjà résolu k problèmes. Deux cas se présentent alors :

  × si $k \in [\![0, n]\!]$, alors l'événement G_n est réalisé si et seulement s'il existe un instant $t' \geqslant t$ où le nombre de problèmes résolus par le groupe A est strictement supérieur au nombre de problèmes résolus par le groupe B
  (on rappelle que t est l'instant où le problème Q_n est résolu par le groupe B)

  Comme l'événement $[T_n = k]$ est réalisé, à l'instant t :

  - le groupe A a résolu k problèmes ($k \in [\![0, n]\!]$)
  - le groupe B a résolu n problèmes.

  Le groupe A a donc $n-k$ problèmes de retard sur le groupe B. L'événement G_n est donc réalisé si et seulement s'il existe $t' \geqslant t$ où le groupe A a résolu $n-k+1$ problèmes de plus que le groupe B, c'est-à-dire si et seulement si l'événement A_{n-k+1} est réalisé.

  $$\boxed{\text{On en déduit : } \mathbb{P}_{[T_n=k]}(G_n) \;=\; \mathbb{P}(A_{n-k+1}) \;=\; a_{n-k+1}.}$$

  × si $k \in [\![n+1, +\infty[\![$, alors l'événement G_n est toujours réalisé. En effet, à l'instant T où le problème Q_n est résolu par le groupe B :

  - le groupe A a résolu k problèmes ($k > n$)
  - le groupe B a résolu n problèmes.

  $$\boxed{\text{On en déduit : } \mathbb{P}_{[T_n=k]}(G_n) = 1.}$$

- On en conclut :

$$\begin{aligned}\mathbb{P}(G_n) &= \sum_{k=0}^{n} \mathbb{P}([T_n = k])\, \mathbb{P}_{[T_n=k]}(G_n) + \sum_{k=n+1}^{+\infty} \mathbb{P}([T_n = k])\, \mathbb{P}_{[T_n=k]}(G_n) \\ &= \sum_{k=0}^{n} \mathbb{P}([T_n = k])\, a_{n+1-k} + \sum_{k=n+1}^{+\infty} \mathbb{P}([T_n = k]) \\ &= \sum_{k=0}^{n} \mathbb{P}([T_n = k])\, a_{n+1-k} + \mathbb{P}([T_n \geqslant n+1])\end{aligned}$$

$$\boxed{\text{Finalement : } \mathbb{P}(G_n) \;=\; \mathbb{P}(\,[T_n \geqslant n+1]\,) + \sum_{k=0}^{n} \mathbb{P}(\,[T_n = k]\,)\, a_{n+1-k}.}$$

□

b) Dans le cas où $p \geqslant \dfrac{1}{2}$, en déduire : $\mathbb{P}(G_n) = 1$.

Démonstration.
Supposons : $p \geqslant \dfrac{1}{2}$.

- Soit $k \in [\![0, n]\!]$. Comme $n - k + 1 \in \mathbb{N}$, d'après le résultat établi en fin de **Partie II** :

$$a_{n+1-k} = 1$$

Commentaire

- Citer les hypothèses d'un théorème avant son utilisation est **indispensable**. Lorsque ce théorème n'est pas un résultat du cours mais un résultat démontré dans un sujet de concours, ce réflexe doit perdurer.
- On utilise par exemple dans cette question un résultat démontré dans la partie précédente. On n'omettra donc en aucun cas de vérifier que l'on est placé dans le bon cadre d'application de cette propriété (ici $r = n + 1 - k \in \mathbb{N}$).

- On obtient, d'après la question précédente :

$$\begin{aligned}
\mathbb{P}(G_n) &= \sum_{k=0}^{n} \mathbb{P}([T_n = k]) \times 1 + \mathbb{P}([T_n \geqslant n+1]) \\
&= \sum_{k=0}^{n} \mathbb{P}([T_n = k]) + \sum_{k=n+1}^{+\infty} \mathbb{P}([T_n = k]) \\
&= \sum_{k=0}^{+\infty} \mathbb{P}([T_n = k]) \\
&= 1 && \textit{(car } \big([T_n = k] \big)_{k \in \mathbb{N}} \textit{ forme un système complet d'événements)}
\end{aligned}$$

$$\boxed{\mathbb{P}(G_n) = 1}$$

□

c) De même lorsque $p < \dfrac{1}{2}$, démontrer :

$$\mathbb{P}(G_n) = 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} \left(p^k q^n - p^{n+1} q^{k-1} \right)$$

Démonstration.
Supposons : $p < \dfrac{1}{2}$.

- Tout d'abord, d'après la question ***14.a)*** :

$$\mathbb{P}(G_n) = \mathbb{P}([T_n \geqslant n+1]) + \sum_{k=0}^{n} \mathbb{P}([T_n = k])\, a_{n+1-k}$$

- De plus :

$$\begin{aligned}
\mathbb{P}([T_n \geqslant n+1]) &= 1 - \mathbb{P}([T_n < n]) \\
&= 1 - \mathbb{P}([T_n \leqslant n]) && \textit{(car } T_n \textit{ est à valeurs entières)} \\
&= 1 - \sum_{k=0}^{n} \mathbb{P}([T_n = k])
\end{aligned}$$

- On obtient :

$$\begin{aligned}
\mathbb{P}(G_n) &= 1 - \sum_{k=0}^{n} \mathbb{P}([T_n = k]) + \sum_{k=0}^{n} \mathbb{P}([T_n = k])\, a_{n+1-k} \\
&= 1 - \sum_{k=0}^{n} \mathbb{P}([T_n = k])\,(1 - a_{n+1-k}) \\
&= 1 - \sum_{k=0}^{n} \mathbb{P}([T_n = k]) \left(1 - \left(\frac{p}{q}\right)^{n+1-k}\right) && \textit{(d'après le résultat de fin de } \textbf{Partie II}, \textit{ car } n+1-k \in \mathbb{N}) \\
&= 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} p^k\, q^n \left(1 - \frac{p^{n+1-k}}{q^{n+1-k}}\right) && \textit{(d'après } \textbf{13.b)}) \\
&= 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} \left(p^k\, q^n - \frac{p^{n+1}}{q^{1-k}}\right)
\end{aligned}$$

$$\boxed{\text{Finalement : } \mathbb{P}(G_n) = 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} \left(p^k\, q^n - p^{n+1}\, q^{k-1}\right).}$$

□

15. *Une meilleure expression de* $\mathbb{P}(G_n)$ *lorsque* $p < \dfrac{1}{2}$

Pour tout $x \in [0,1]$ et $n \in \mathbb{N}^*$, on pose :

$$u_n(x) = (1-x)^n \sum_{k=0}^{n} \binom{n+k-1}{k} x^k$$

a) Vérifier que pour tout $n \in \mathbb{N}^*$: $\mathbb{P}(G_n) = 1 - u_n(p) + \dfrac{p}{q}\, u_n(q)$.

Démonstration.
Soit $n \in \mathbb{N}^*$.

$$\begin{aligned}
&1 - u_n(p) + \frac{p}{q}\, u_n(q) \\
=\ & 1 - (1-p)^n \sum_{k=0}^{n} \binom{n+k-1}{k} p^k + \frac{p}{q}\,(1-q)^n \sum_{k=0}^{n} \binom{n+k-1}{k} q^k \\
=\ & 1 - q^n \sum_{k=0}^{n} \binom{n+k-1}{k} p^k + \frac{p}{q}\, p^n \sum_{k=0}^{n} \binom{n+k-1}{k} q^k \\
=\ & 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} p^k\, q^n + \sum_{k=0}^{n} \binom{n+k-1}{k} p^{n+1}\, q^{k-1} \\
=\ & 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} \left(p^k\, q^n - p^{n+1}\, q^{k-1}\right) \\
=\ & \mathbb{P}(G_n) && \textit{d'après } \textbf{14.c)}
\end{aligned}$$

$$\boxed{\forall n \in \mathbb{N}^*,\ \mathbb{P}(G_n) = 1 - u_n(p) + \frac{p}{q}\, u_n(q)}$$

□

b) Pour tout $x \in [0,1]$ et $n \in \mathbb{N}^*$, établir la relation :

$$u_{n+1}(x) \;=\; u_n(x) + (1-x)^n\, x^{n+1} \left(\binom{2n}{n+1} - \binom{2n+1}{n+1}\, x \right)$$

Démonstration.
Soit $x \in [0,1]$. Soit $n \in \mathbb{N}^*$.

- Tout d'abord :

$$\begin{aligned}
& u_{n+1}(x) \\
=\;& (1-x)^{n+1} \sum_{k=0}^{n+1} \binom{(n+\cancel{1})+k-\cancel{1}}{k}\, x^k \\
=\;& (1-x)\,(1-x)^n \sum_{k=0}^{n+1} \binom{n+k}{k}\, x^k \\
=\;& (1-x)^n \sum_{k=0}^{n+1} \binom{n+k}{k}\, x^k - x\,(1-x)^n \sum_{k=0}^{n+1} \binom{n+k}{k}\, x^k \\
=\;& (1-x)^n \sum_{k=0}^{n+1} \binom{n+k}{k}\, x^k - x\,(1-x)^n \left(\sum_{k=0}^{n} \binom{n+k}{k}\, x^k + \binom{2n+1}{n+1}\, x^{n+1} \right) \\
=\;& (1-x)^n \sum_{k=0}^{n+1} \binom{n+k}{k}\, x^k - x\,(1-x)^n \sum_{k=0}^{n} \binom{n+k}{k}\, x^k - x^{n+1}\,(1-x)^n \binom{2n+1}{n+1}\, x \qquad (*)
\end{aligned}$$

- Simplifions légèrement le 2$^{\text{ème}}$ terme de la somme $(*)$ ci-dessus.

$$\begin{aligned}
x\,(1-x)^n \sum_{k=0}^{n} \binom{n+k}{k}\, x^k \;&=\; (1-x)^n \sum_{k=0}^{n} \binom{n+k}{k}\, x^{k+1} \\
&=\; (1-x)^n \sum_{k=1}^{n+1} \binom{n+k-1}{k-1}\, x^k \qquad \textit{(par décalage d'indice)}
\end{aligned}$$

- Rassemblons maintenant les 2 premiers termes de $(*)$.

$$\begin{aligned}
& (1-x)^n \sum_{k=0}^{n+1} \binom{n+k}{k}\, x^k - (1-x)^n \sum_{k=1}^{n+1} \binom{n+k-1}{k-1}\, x^k \\
=\;& (1-x)^n \left(\sum_{k=0}^{n+1} \binom{n+k}{k}\, x^k - \sum_{k=1}^{n+1} \binom{n+k-1}{k-1}\, x^k \right) \\
=\;& (1-x)^n \left(\binom{n}{0}\, x^0 + \sum_{k=1}^{n+1} \binom{n+k}{k}\, x^k - \sum_{k=1}^{n+1} \binom{n+k-1}{k-1}\, x^k \right) \\
=\;& (1-x)^n \left(1 + \sum_{k=1}^{n+1} \left(\binom{n+k}{k} - \binom{n+k-1}{k-1} \right) x^k \right) \\
=\;& (1-x)^n \left(1 + \sum_{k=1}^{n+1} \binom{n+k-1}{k}\, x^k \right) \qquad \textit{(par triangle de Pascal)} \\
=\;& (1-x)^n \sum_{k=0}^{n+1} \binom{n+k-1}{k}\, x^k
\end{aligned}$$

- On en déduit :

$$\begin{aligned}
& u_{n+1}(x) \\
= \;& (1-x)^n \sum_{k=0}^{n+1} \binom{n+k-1}{k} x^k - x^{n+1}(1-x)^n \binom{2n+1}{n+1} x \\
= \;& (1-x)^n \left(\sum_{k=0}^{n} \binom{n+k-1}{k} x^k + \binom{2n}{n+1} x^{n+1} \right) - x^{n+1}(1-x)^n \binom{2n+1}{n+1} x \\
= \;& (1-x)^n \sum_{k=0}^{n} \binom{n+k-1}{k} x^k + x^{n+1}(1-x)^n \binom{2n}{n+1} - x^{n+1}(1-x)^n \binom{2n+1}{n+1} x \\
= \;& u_n(x) + x^{n+1}(1-x)^n \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} x \right)
\end{aligned}$$

$$\boxed{\forall x \in [0,1],\ \forall n \in \mathbb{N}^*,\ u_{n+1}(x) = u_n(x) + x^{n+1}(1-x)^n \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} x \right)}$$

□

c) En déduire, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(G_{n+1}) \;=\; \mathbb{P}(G_n) - \left(1 - \frac{p}{q}\right) (pq)^{n+1} \binom{2n+1}{n+1}$$

Démonstration.
Soit $n \in \mathbb{N}^*$.

$$\begin{aligned}
& \mathbb{P}(G_{n+1}) \\
= \;& 1 - u_{n+1}(p) + \frac{p}{q}\, u_{n+1}(q) && \textit{(d'après } \textbf{\textit{15.a)}}\textit{)} \\
= \;& 1 - \left(u_n(p) + p^{n+1} q^n \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} p \right) \right) + \frac{p}{q} \left(u_n(q) + q^{n+1} p^n \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} q \right) \right) \\
& && \textit{(d'après la question précédente)} \\
= \;& \left(1 - u_n(p) + \frac{p}{q}\, u_n(q) \right) - p^{n+1} q^n \left(\cancel{\binom{2n}{n+1}} - \binom{2n+1}{n+1} p \right) + p^{n+1} q^n \left(\cancel{\binom{2n}{n+1}} - \binom{2n+1}{n+1} q \right) \\
= \;& \mathbb{P}(G_n) - p^{n+1} q^n \binom{2n+1}{n+1} (q-p) && \textit{(d'après } \textbf{\textit{15.a)}}\textit{)} \\
= \;& \mathbb{P}(G_n) - p^{n+1} q^{n+1} \frac{1}{q} \binom{2n+1}{n+1} (q-p) \\
= \;& \mathbb{P}(G_n) - (p\,q)^{n+1} \binom{2n+1}{n+1} \left(1 - \frac{p}{q}\right)
\end{aligned}$$

$$\boxed{\forall n \in \mathbb{N}^*,\ \mathbb{P}(G_{n+1}) \;=\; \mathbb{P}(G_n) - \left(1 - \frac{p}{q}\right) (pq)^{n+1} \binom{2n+1}{n+1}}$$

□

d) Montrer par récurrence, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(G_n) \;=\; \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k$$

Démonstration.
Démontrons par récurrence : $\forall n \in \mathbb{N}^*, \mathcal{P}(n)$ où $\mathcal{P}(n) : \mathbb{P}(G_n) \;=\; \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k$.

▸ **Initialisation** :

- D'une part, d'après ***14.c)*** :

$$\begin{aligned}
\mathbb{P}(G_1) &= 1 - \sum_{k=0}^{1} \binom{\cancel{1}+k-\cancel{1}}{k} \left(p^k\, q - p^2\, q^{k-1}\right) \\
&= 1 - \left(\left(p^0\, q - \frac{p^2}{q}\right) + \left(p\, q - p^2\, q^0\right)\right) \\
&= (1-q) + \frac{p^2}{q} - p\, q + p^2 \\
&= p + \frac{p^2}{q} - p\, q + p^2
\end{aligned}$$

- D'autre part :

$$\begin{aligned}
\frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{1} \binom{2k-1}{k} (p\, q)^k &= \frac{p}{q} - \left(1 - \frac{p}{q}\right) \binom{1}{1} (p\, q)^1 \\
&= \frac{p}{q} - p\, q \left(1 - \frac{p}{q}\right) \\
&= \frac{p}{q} - p\, q + p^2
\end{aligned}$$

- Vérifions alors : $p + \frac{p^2}{q} - p\, q + p^2 = \frac{p}{q} - p\, q + p^2$.

$$\begin{aligned}
p + \frac{p^2}{q} - \cancel{p\, q} + \cancel{p^2} - \left(\frac{p}{q} - \cancel{p\, q} + \cancel{p^2}\right) &= p + \frac{p^2}{q} - \frac{p}{q} \\
&= p\left(1 + \frac{p}{q} - \frac{1}{q}\right) \\
&= p\ \frac{q+p-1}{q} \\
&= p\ \frac{1-1}{q} \;=\; 0 \qquad \textit{(car } p+q=1\textit{)}
\end{aligned}$$

On a ainsi bien démontré :

$$\mathbb{P}(G_1) \;=\; \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{1} \binom{2k-1}{k} (p\, q)^k$$

D'où $\mathcal{P}(1)$.

▸ **Hérédité** : soit $n \in \mathbb{N}^*$.
Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$ (*i.e.* $\mathbb{P}(G_{n+1}) \; = \; \dfrac{p}{q} - \left(1 - \dfrac{p}{q}\right) \sum\limits_{k=1}^{n+1} \dbinom{2k-1}{k} \, (pq)^k$).

$$\begin{array}{rll} & \mathbb{P}(G_{n+1}) & \\ = & \mathbb{P}(G_n) - \left(1 - \dfrac{p}{q}\right) (p\,q)^{n+1} \dbinom{2n+1}{n+1} & \textit{(d'après \textbf{15.c}))} \\ = & \left(\dfrac{p}{q} - \left(1 - \dfrac{p}{q}\right) \sum\limits_{k=1}^{n} \dbinom{2k-1}{k} \, (p\,q)^k\right) - \left(1 - \dfrac{p}{q}\right) (p\,q)^{n+1} \dbinom{2n+1}{n+1} & \textit{(par hypothèse de récurrence)} \\ = & \dfrac{p}{q} - \left(1 - \dfrac{p}{q}\right) \left(\sum\limits_{k=1}^{n} \dbinom{2k-1}{k} \, (p\,q)^k + \dbinom{2n+1}{n+1} \, (p\,q)^{n+1}\right) & \\ = & \dfrac{p}{q} - \left(1 - \dfrac{p}{q}\right) \sum\limits_{k=1}^{n+1} \dbinom{2k-1}{k} \, (pq)^k & \end{array}$$

D'où $\mathcal{P}(n+1)$.

$$\boxed{\text{Par principe de récurrence : } \forall n \in \mathbb{N}^*, \, \mathbb{P}(G_n) \; = \; \dfrac{p}{q} - \left(1 - \dfrac{p}{q}\right) \sum\limits_{k=1}^{n} \dbinom{2k-1}{k} \, (pq)^k.}$$

□

16. *Application à la sécurisation des transactions*
Connaissant $p < \dfrac{1}{2}$, on cherche à limiter le risque que la stratégie mise en place par le groupe de mineurs A réussisse.

a) Après avoir établi la formule $\dbinom{n}{k} = \dfrac{n}{k} \dbinom{n-1}{k-1}$ lorsque $k \in [\![1, n]\!]$, écrire une fonction **Python** qui calcule les coefficients binomiaux.

Démonstration.
Soit $k \in [\![1, n]\!]$.

- Tout d'abord :

$$k \dbinom{n}{k} = k \; \frac{n!}{k! \; (n-k)!} = \frac{n!}{(k-1)! \; (n-k)!}$$

- Par ailleurs :

$$n \dbinom{n-1}{k-1} = n \; \frac{(n-1)!}{(k-1)! \; ((n-\cancel{1}) - (k-\cancel{1}))!} = \frac{n!}{(k-1)! \; (n-k)!}$$

Ainsi : $k \dbinom{n}{k} = n \dbinom{n-1}{k-1}$.

$$\boxed{\text{D'où : } \forall k \in [\![1, n]\!], \; \dbinom{n}{k} = \dfrac{n}{k} \dbinom{n-1}{k-1}.}$$

Commentaire

La relation sur les coefficients binomiaux peut aussi se faire par dénombrement.
Pour ce faire, on considère un ensemble E à n éléments.
(on peut penser à une pièce qui contient n individus)
On souhaite alors construire une partie P à k éléments de cet ensemble contenant un élément distingué *(on peut penser à choisir dans la pièce un groupe de k individus dans lequel figure un représentant de ces individus).*
Pour ce faire, on peut procéder de deux manières :

1) On choisit d'abord la partie à k éléments de E : $\binom{n}{k}$ possibilités.
On distingue ensuite un élément de cet ensemble P : $\binom{k}{1} = k$ possibilités.
(on choisit d'abord les k individus et on élit ensuite un représentant de ces individus)
Ainsi, il y a $k\,\binom{n}{k}$ manières de construire P.

2) On choisit d'abord, dans E, l'élément à distinguer : $\binom{n}{1} = n$ possibilités.
On choisit ensuite $k-1$ éléments dans E qui, pour former P, en y ajoutant l'élément précédent : $\binom{n-1}{k-1}$ possibilités.
(on choisit d'abord le représentant puis on lui adjoint un groupe de $k-1$ individus)
Ainsi, il y a $n\,\binom{n-1}{k-1}$ manières de construire P.

On retrouve ainsi le résultat.

- En itérant la formule précédente, on obtient :

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} = \frac{n}{k}\,\frac{n-1}{k-1}\binom{n-2}{k-2} = \cdots = \frac{n\,(n-1)\,\cdots\,(n-k+1)}{k\,(k-1)\,\cdots\,1}$$

(cette formule se démontre rigoureusement par récurrence)

- On propose alors la fonction **Python** suivante.

```
1 def coeffBin(k, n):
2     c = 1
3     for i in range(1, k+1):
4         c = c * (n-i+1) / i
5     return c
```

Détaillons les éléments de ce script.

- **Début de la fonction**
On commence par préciser la structure de la fonction :
  × cette fonction se nomme `coeffBin`,
  × elle prend en paramètre les variables `k` et `n`.

```
1 def coeffBin(k, n):
```

On initialise ensuite la variable `c` à 1 (choix naturel d'initialisation lorsqu'on souhaite coder un produit puisque 1 est l'élément neutre de l'opérateur produit).

```
2     c = 1
```

- **Structure itérative**
Les lignes 3 à 4 consistent à mettre à jour la variable `c` pour qu'elle contienne la quantité $\binom{n}{k}$.

Or, d'après ce qui précède :

$$\binom{n}{k} = \frac{n\,(n-1)\,\cdots\,(n-k+1)}{k\,(k-1)\,\cdots\,1} = \frac{\prod\limits_{i=1}^{k}(n-i+1)}{\prod\limits_{i=1}^{k} i} = \prod_{i=1}^{k} \frac{n-i+1}{i}$$

Pour cela, on utilise alors une structure conditionnelle (boucle `for`) :

```
    for i in range(1, k+1):
        c = c * (n-i+1) / i
```

- **Fin de la fonction**
  À l'issue de cette boucle, la variable `c` contient la quantité $\prod\limits_{i=1}^{k} \dfrac{n-i+1}{i} = \dbinom{n}{k}$.

Commentaire

- Afin de permettre une bonne compréhension des mécanismes en jeu, on a détaillé la réponse à cette question. Cependant, proposer un programme **Python** correct démontre la bonne de ces mécanismes et permet certainement d'obtenir la majorité des points alloués à cette question. On procèdera de même dans les autres questions **Python**.
- Comme expliqué plus haut, on initialise `c` à 1 puisque cette variable doit contenir un produit, et que le réel 1 est l'élément neutre pour l'opérateur produit.
  On rappelle qu'on procède de même avec l'initialisation d'une somme stockée dans une variable `S` : on initialise la variable `S` à 0 car le réel 0 est l'élément neutre pour l'opérateur de sommation.
- On remarque que le programme proposé permet bien d'obtenir : $\dbinom{n}{0} = 1$.
  En effet, si $\texttt{k} = 0$, alors :
  1) la variable `c` est initialisée à 1,
  2) la boucle qui suit n'est pas effectuée puisque la matrice `1:0` est une matrice vide,
  3) la fonction renvoie donc bien 1 lorsque `k` vaut 0.
- On pouvait exploiter plus directement la relation : $\dbinom{n}{k} = \dfrac{n}{k}\dbinom{n-1}{k-1}$.
  On obtient la fonction **Python** suivante :

```
def coeffBinRec(k, n):
if k == 0:
c = 1
else:
c = (n/k) * coeffBinRec(k-1, n-1)
return c
```

  On remarque que la définition de la fonction `coeffBinRec` fait appel à elle-même. On dit que la fonction `coeffBinRec` est définie de manière *récursive*. Cette manière de coder est ici rendue naturelle par la formule démontrée juste avant.

Commentaire

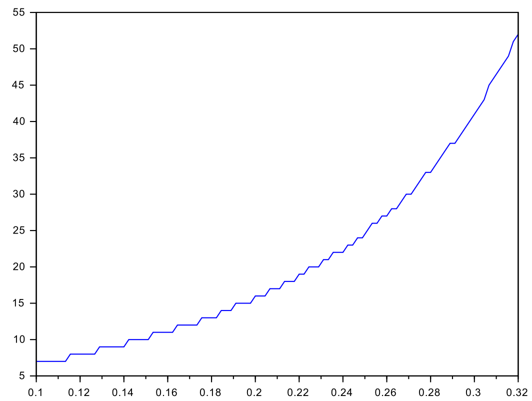
- Par exemple, lorsqu'on effectue l'appel `coeffBinRec(2,3)` (pour obtenir la valeur de $\binom{3}{2}$), le calcul s'effectue de la façon suivante :

$$\begin{aligned}\texttt{coeffBinRec}(2,3) &= \frac{3}{2} \times \texttt{coeffBinRec}(2,1) \\ &= \frac{3}{2} \times \frac{2}{1} \times \texttt{coeffBinRec}(1,0) \\ &= \frac{3}{2} \times \frac{2}{1} \times 1 \;=\; \frac{3 \times 2}{2 \times 1} \;=\; \binom{3}{2}\end{aligned}$$

Remarquons que le calcul est certain d'aboutir puisque l'appel `coeffBinRec(k,n)` nécessite les appels de `coeffBinRec(k-1, n-1)`, puis `coeffBinRec(k-2, n-2)`, ..., puis `coeffBinRec(1, n-(k-2))`, et enfin `coeffBinRec(0, n-(k-1))` (dont on connaît la valeur).

b) Écrire un script **Python** qui détermine n_p, le plus petit entier n tel que $\mathbb{P}(G_n) \leqslant \varepsilon$ pour $p < \frac{1}{2}$ et $\varepsilon > 0$ saisis au clavier par l'utilisateur.

NB : Pour $\varepsilon = 10^{-4} = 0,1\%$ et p variant entre 10% et 32%, on obtient pour la représentation de n_p en fonction de p :



Démonstration.
On rappelle le résultat suivant, obtenu à la question ***15.d)*** pour le cas $p < \frac{1}{2}$:

$$\mathbb{P}(G_n) \;=\; \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k$$

Ainsi :

$$\mathbb{P}(G_n) \leqslant \varepsilon \iff \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k \geqslant \frac{\frac{p}{q} - \varepsilon}{1 - \frac{p}{q}}$$

On propose alors le script **Python** suivant, qui repose sur le calcul de la somme $\sum\limits_{k=1}^{n} \binom{2k-1}{k} (pq)^k$.

```
p = float(input('p = '))
eps = float(input('eps = '))
q = 1 - p
borne = (p/q - eps)/(1 - p/q)
S = 0
n = 0
while S < borne:
    n = n + 1
    S = S + coeffBin(n, 2*n-1) * (p*q)**n
print(n)
```

□

HEC 2008 (Exercice) - fonction de deux variables, méthode des moindres carrés, droite de régression linéaire

Étant donné un entier n supérieur ou égal à 2, on considère un nuage de n points du plan, c'est-à-dire un n-uplet $\big((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\big)$ d'éléments de $\mathbb{R}^2$. On suppose que les réels x_1, x_2, …, x_n (resp. y_1, y_2, …, y_n) ne sont pas tous égaux.
On appelle moyenne arithmétique $\overline{x}$ et écart-type σ_x du n-uplet $x = (x_1, \dots, x_n)$, les réels suivants :

$$\overline{x} = \frac{1}{n} \sum_{k=1}^{n} x_k \qquad \text{et} \qquad \sigma_x = \sqrt{\frac{1}{n} \sum_{k=1}^{n} (x_k - \overline{x})^2}$$

On définit de même la moyenne arithmétique $\overline{y}$ et l'écart-type σ_y du n-uplet $y = (y_1, \dots, y_n)$.
La covariance $\text{cov}(x, y)$ et le coefficient de corrélation linéaire $r(x, y)$ du couple (x, y) sont donnés par :

$$\text{cov}(x, y) = \frac{1}{n} \sum_{k=1}^{n} (x_k - \overline{x})(y_k - \overline{y}) \qquad \text{et} \qquad r(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \times \sigma_y}$$

Soit f la fonction définie sur $\mathbb{R}^2$ à valeurs réelles qui, à tout couple (a, b) de $\mathbb{R}^2$, associe le réel $f(a, b)$ tel que :

$$f(a, b) = \sum_{k=1}^{n} (a\, x_k + b - y_k)^2$$

1. Justifier que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.

Démonstration.

- Soit $(a, b) \in \mathbb{R}^2$.

$$\begin{aligned} & f(a, b) \\ = \ & \sum_{k=1}^{n} (a\, x_k + b - y_k)^2 \\ = \ & \sum_{k=1}^{n} \big(a^2 (x_k)^2 + b^2 + (y_k)^2 + 2\, a\, b\, x_k - 2\, a\, x_k\, y_k - 2\, b\, y_k\big) \\ = \ & \left(\sum_{k=1}^{n} (x_k)^2\right) a^2 + n\, b^2 + 2 \left(\sum_{k=1}^{n} x_k\right) a\, b - 2 \left(\sum_{k=1}^{n} x_k\, y_k\right) a - 2 \left(\sum_{k=1}^{n} y_k\right) b + \sum_{k=1}^{n} (y_k)^2 \end{aligned}$$

- La fonction f est donc polynomiale.

$$\boxed{\text{On en déduit que la fonction } f \text{ est de classe } \mathcal{C}^2 \text{ sur } \mathbb{R}^2.}$$

Commentaire

On peut aussi remarquer que f est une somme de carrés de fonctions polynomiales et donc est, à ce titre, une fonction polynomiale. Les calculs ci-dessus sont en réalité superflus si on a en tête que l'espace des fonctions polynomiales est stable par somme et par produit.

□

2. a) Écrire le système d'équations (S) permettant de déterminer les points critiques de f.

Démonstration.

- Comme f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$, elle admet des dérivées partielles d'ordre 1 (et 2) sur $\mathbb{R}^2$. Déterminons les.
- Soit $(a,b) \in \mathbb{R}^2$.
  - × Tout d'abord :

$$\begin{array}{rcll} \partial_1(f)(a,b) & = & \sum\limits_{k=1}^{n} \big(2\,x_k\,(a\,x_k + b - y_k)\big) & \\ & = & 2\left(a \sum\limits_{k=1}^{n} (x_k)^2 + b \sum\limits_{k=1}^{n} x_k - \sum\limits_{k=1}^{n} x_k\,y_k\right) & \\ & = & 2\left(a \sum\limits_{k=1}^{n} (x_k)^2 + b \times (n\,\bar{x}) - \sum\limits_{k=1}^{n} x_k\,y_k\right) & \textit{(par définition de } \bar{x}\textit{)} \end{array}$$

Exprimons maintenant $\sum\limits_{k=1}^{n} (x_k)^2$ et $\sum\limits_{k=1}^{n} x_k\,y_k$ en fonction de $\bar{x}$, $\bar{y}$, σ_x et $\mathrm{cov}(x,y)$ pour simplifier l'expression de $\partial_1(f)(a,b)$.

- Comme $\sigma_x = \sqrt{\dfrac{1}{n} \sum\limits_{k=1}^{n} (x_k - \bar{x})^2}$, on obtient :

$$\begin{array}{rcl} \sigma_x^2 & = & \dfrac{1}{n} \sum\limits_{k=1}^{n} (x_k - \bar{x})^2 \\ & = & \dfrac{1}{n} \sum\limits_{k=1}^{n} \big((x_k)^2 - 2\,\bar{x}\,x_k + \bar{x}^2\big) \\ & = & \dfrac{1}{n} \left(\sum\limits_{k=1}^{n} (x_k)^2 - 2\,\bar{x} \sum\limits_{k=1}^{n} x_k + n\,\bar{x}^2\right) \\ & = & \dfrac{1}{n} \sum\limits_{k=1}^{n} (x_k)^2 - 2\,\bar{x} \times \left(\dfrac{1}{n} \sum\limits_{k=1}^{n} x_k\right) + \bar{x}^2 \\ & = & \dfrac{1}{n} \sum\limits_{k=1}^{n} (x_k)^2 - 2\,\bar{x}^2 + \bar{x}^2 \\ & = & \dfrac{1}{n} \sum\limits_{k=1}^{n} (x_k)^2 - \bar{x}^2 \end{array}$$

Ainsi : $\dfrac{1}{n} \sum\limits_{k=1}^{n} (x_k)^2 = \sigma_x^2 + \bar{x}^2$.

$$\boxed{\text{D'où : } \sum\limits_{k=1}^{n} (x_k)^2 = n\big(\sigma_x^2 + \bar{x}^2\big).}$$

- De plus :

$$\begin{aligned}
\operatorname{cov}(x,y) &= \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})\,(y_k - \bar{y}) \\
&= \frac{1}{n} \sum_{k=1}^{n} (x_k\, y_k - \bar{y}\, x_k - \bar{x}\, y_k + \bar{x}\, \bar{y}) \\
&= \frac{1}{n} \left(\sum_{k=1}^{n} x_k\, y_k - \bar{y} \sum_{k=1}^{n} x_k - \bar{x} \sum_{k=1}^{n} y_k + n\, \bar{x}\, \bar{y} \right) \\
&= \frac{1}{n} \sum_{k=1}^{n} x_k\, y_k - \bar{y} \left(\frac{1}{n} \sum_{k=1}^{n} x_k \right) - \bar{x} \left(\frac{1}{n} \sum_{k=1}^{n} y_k \right) + \bar{x}\, \bar{y} \\
&= \frac{1}{n} \sum_{k=1}^{n} x_k\, y_k - \bar{y}\, \bar{x} - \cancel{\bar{x}\, \bar{y}} + \cancel{\bar{x}\, \bar{y}} \\
&= \frac{1}{n} \sum_{k=1}^{n} x_k\, y_k - \bar{x}\, \bar{y}
\end{aligned}$$

Ainsi : $\displaystyle\frac{1}{n} \sum_{k=1}^{n} x_k\, y_k = \operatorname{cov}(x,y) + \bar{x}\, \bar{y}$.

$$\boxed{\text{D'où : } \sum_{k=1}^{n} x_k\, y_k = n\big(\operatorname{cov}(x,y) + \bar{x}\, \bar{y}\big).}$$

On en déduit :

$$\begin{aligned}
\partial_1(f)(a,b) &= 2 \left(a \sum_{k=1}^{n} (x_k)^2 + n\, \bar{x}\, b - \sum_{k=1}^{n} x_k\, y_k \right) \\
&= 2\Big(n\ \left(\sigma_x^2 + \bar{x}^2\right)\ a + n\, \bar{x}\, b -,\ \big(\operatorname{cov}(x,y) + \bar{x}\, \bar{y}\big)\Big)
\end{aligned}$$

$$\boxed{\forall (a,b) \in \mathbb{R}^2,\ \partial_1(f)(a,b) = 2\, n \Big(\left(\sigma_x^2 + \bar{x}^2\right)\ a + \bar{x}\, b - \operatorname{cov}(x,y) - \bar{x}\, \bar{y} \Big)}$$

× Soit $(a,b) \in \mathbb{R}^2$.

$$\begin{aligned}
\partial_2(f)(a,b) &= \sum_{k=1}^{n} \big(2 \times 1 \times (a\, x_k + b - y_k)\big) \\
&= 2 \left(a \sum_{k=1}^{n} x_k + n\, b - \sum_{k=1}^{n} y_k \right) \\
&= 2\big(a \times (n\, \bar{x}) + n\, b - n\, \bar{y}\big) && \textit{(par définition de } \bar{x} \textit{ et } \bar{y}\textit{)} \\
&= 2\, n\ (a\, \bar{x} + b - \bar{y})
\end{aligned}$$

$$\boxed{\forall (a,b) \in \mathbb{R}^2,\ \partial_2(f)(a,b) = 2\, n\ (a\, \bar{x} + b - \bar{y})}$$

- Soit $(a,b) \in \mathbb{R}^2$.

$$\begin{array}{rcl} (a,b) \text{ est un point critique de } f & \Leftrightarrow & \nabla(f)(a,b) = 0_{\mathscr{M}_{2,1}(\mathbb{R})} \\ & \Leftrightarrow & \left\{ \begin{array}{rcl} \partial_1(f)(a,b) & = & 0 \\ \partial_2(f)(a,b) & = & 0 \end{array} \right. \end{array}$$

Ainsi, d'après la question précédente :

$$\begin{array}{rcl} \begin{array}{c} (a,b) \text{ est un point} \\ \text{critique de } f \end{array} & \Leftrightarrow & \left\{ \begin{array}{l} 2\,n\,\Big(\big(\sigma_x^2 + \bar{x}^2\big)\ a + \bar{x}\,b - \mathrm{cov}(x,y) - \bar{x}\,\bar{y} \Big) \ = \ 0 \\ 2\,n\ (a\,\bar{x} + b - \bar{y}) \ = \ 0 \end{array} \right. \\ & \Leftrightarrow & \left\{ \begin{array}{l} \big(\sigma_x^2 + \bar{x}^2\big)\ a + \bar{x}\,b - \mathrm{cov}(x,y) - \bar{x}\,\bar{y} \ = \ 0 \\ a\,\bar{x} + b - \bar{y} \ = \ 0 \end{array} \right. \end{array}$$

$$\boxed{\text{Le système } (S) \text{ souhaité est le système : } \left\{ \begin{array}{l} \big(\sigma_x^2 + \bar{x}^2\big)\ a + \bar{x}\,b - \mathrm{cov}(x,y) - \bar{x}\,\bar{y} \ = \ 0 \\ a\,\bar{x} + b - \bar{y} \ = \ 0 \end{array} \right. .}$$

□

b) Résoudre le système (S). En déduire que f admet un unique point critique $(\hat{a}, \hat{b})$ que l'on exprimera en fonction de $\bar{x}$, $\bar{y}$, σ_x^2 et $\mathrm{cov}(x,y)$.

Démonstration.
Soit $(a,b) \in \mathbb{R}^2$.

$$\begin{array}{rcl} \begin{array}{c} (a,b) \text{ est un point} \\ \text{critique de } f \end{array} & \Leftrightarrow & \left\{ \begin{array}{l} \big(\sigma_x^2 + \bar{x}^2\big)\ a + \bar{x}\,b - \mathrm{cov}(x,y) - \bar{x}\,\bar{y} \ = \ 0 \\ a\,\bar{x} + b - \bar{y} \ = \ 0 \end{array} \right. \\ & \Leftrightarrow & \left\{ \begin{array}{l} \big(\sigma_x^2 + \bar{x}^2\big)\ a + \bar{x}\,b \ = \ \mathrm{cov}(x,y) + \bar{x}\,\bar{y} \\ b \ = \ \bar{y} - a\,\bar{x} \end{array} \right. \\ & \Leftrightarrow & \left\{ \begin{array}{l} \big(\sigma_x^2 + \bar{x}^2\big)\ a + \bar{x}\,(\bar{y} - a\,\bar{x}) \ = \ \mathrm{cov}(x,y) + \bar{x}\,\bar{y} \\ b \ = \ \bar{y} - a\,\bar{x} \end{array} \right. \\ & \Leftrightarrow & \left\{ \begin{array}{l} \big(\sigma_x^2 + \cancel{\bar{x}^2} - \cancel{\bar{x}^2}\big)\ a + \cancel{\bar{x}\bar{y}} \ = \ \mathrm{cov}(x,y) + \cancel{\bar{x}\bar{y}} \\ b \ = \ \bar{y} - a\,\bar{x} \end{array} \right. \\ & \Leftrightarrow & \left\{ \begin{array}{l} a \ = \ \dfrac{\mathrm{cov}(x,y)}{\sigma_x^2} \\ b \ = \ \bar{y} - a\,\bar{x} \end{array} \right. \\ & \Leftrightarrow & \left\{ \begin{array}{l} a \ = \ \dfrac{\mathrm{cov}(x,y)}{\sigma_x^2} \\ b \ = \ \bar{y} - \dfrac{\mathrm{cov}(x,y)}{\sigma_x^2}\ \bar{x} \end{array} \right. \end{array}$$

$$\boxed{\text{On en déduit que } f \text{ admet un unique point critique } (\hat{a}, \hat{b}) = \left(\frac{\mathrm{cov}(x,y)}{\sigma_x^2}\ ,\ \bar{y} - \frac{\mathrm{cov}(x,y)}{\sigma_x^2}\ \bar{x} \right).}$$

Commentaire

- La difficulté de la recherche de points critiques réside dans le fait qu'il n'existe pas de méthode générale pour résoudre l'équation $\nabla(f)(a,b) = 0_{\mathscr{M}_{2,1}(\mathbb{R})}$.
On est donc confronté à une question bien plus complexe qu'une résolution de système d'équations linéaires (que l'on résout aisément à l'aide de la méthode du pivot de Gauss).
- Lors de la recherche de points critiques, on doit faire appel à des méthodes ad hoc. Ici, on fait apparaître une équation du type :
$$b = \psi(a)$$
En injectant cette égalité dans la seconde équation, on obtient une nouvelle équation qui ne dépend plus que d'une variable et qu'il est donc plus simple de résoudre.
C'est la stratégie qu'on a adoptée ci-dessus.

□

c) Montrer que ce point critique correspond à un minimum local de f.

Démonstration.

- La fonction f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$. Elle admet donc des dérivées partielles secondes sur $\mathbb{R}^2$.
- Soit $(x,y) \in \mathbb{R}^2$.
  × Tout d'abord :
  $$\partial^2_{1,1}(f)(a,b) \ = \ 2\,n\;(\sigma_x^2 + \bar{x}^2)$$
  × Ensuite :
  $$\partial^2_{1,2}(f)(a,b) \ = \ \partial^2_{2,1}(f)(a,b) \ = \ 2\,n\,\bar{x}$$
  La première égalité est obtenue en vertu du théorème de Schwarz puisque la fonction f est de classe $\mathcal{C}^2$ sur l'**ouvert** $\mathbb{R}^2$.
  × Enfin :
  $$\partial^2_{2,2}(f)(a,b) \ = \ 2\,n$$

$$\boxed{\begin{aligned} \forall (a,b) \in \mathbb{R}^2, \quad \partial^2_{1,1}(f)(a,b) \ &= \ 2\,n\;(\sigma_x^2 + \bar{x}^2) \\ \partial^2_{1,2}(f)(a,b) \ &= \ \partial^2_{2,1}(f)(a,b) \ = \ 2\,n\,\bar{x} \\ \partial^2_{2,2}(f)(a,b) \ &= \ 2\,n \end{aligned}}$$

Commentaire

- Il faut penser à utiliser le théorème de Schwarz dès que la fonction à deux variables considérée est de classe $\mathcal{C}^2$ sur un ouvert $U \subset \mathbb{R}^2$.
- Ici, le calcul de $\partial^2_{1,2}(f)(a,b)$ et $\partial^2_{2,1}(f)(a,b)$ est aisé. Il faut alors concevoir le résultat du théorème de Schwarz comme une mesure de vérification : en dérivant par rapport à la $1^{\text{ère}}$ variable puis par rapport à la $2^{\text{ème}}$, on doit obtenir le même résultat que dans l'ordre inverse.

- Soit $(a,b) \in \mathbb{R}^2$. On rappelle que la matrice hessienne de f en (a,b) est :

$$\nabla^2(f)(a,b) \;=\; \begin{pmatrix} \partial^2_{1,1}(f)(a,b) & \partial^2_{1,2}(f)(a,b) \\ \partial^2_{2,1}(f)(a,b) & \partial^2_{2,2}(f)(a,b) \end{pmatrix} \;=\; \begin{pmatrix} 2\,n\,(\sigma_x^2 + \bar{x}^2) & 2\,n\,\bar{x} \\ 2\,n\,\bar{x} & 2\,n \end{pmatrix}$$

$$\boxed{\text{On en déduit : } H \;=\; \nabla^2(f)(\hat{a},\hat{b}) \;=\; \begin{pmatrix} 2\,n\,(\sigma_x^2 + \bar{x}^2) & 2\,n\,\bar{x} \\ 2\,n\,\bar{x} & 2\,n \end{pmatrix}.}$$

- Soit $\lambda \in \mathbb{R}$.

$$\begin{aligned} \det(H - \lambda\, I_2) &= \det\left(\begin{pmatrix} 2\,n\;(\sigma_x^2 + \bar{x}^2) - \lambda & 2\,n\,\bar{x} \\ 2\,n\,\bar{x} & 2\,n - \lambda \end{pmatrix}\right) \\ &= \big(2\,n\,(\sigma_x^2 + \bar{x}^2) - \lambda\big)\;(2\,n - \lambda) - (2\,n\,\bar{x})^2 \\ &= \lambda^2 - \big(2\,n\,(\sigma_x^2 + \bar{x}^2) + 2\,n\big)\,\lambda + 4\,n^2\,(\sigma_x^2 + \cancel{\bar{x}^2}) - \cancel{4\,n^2\,\bar{x}^2} \\ &= \lambda^2 - 2\,n\,(\sigma_x^2 + \bar{x}^2 + 1)\,\lambda + 4\,n^2\,\sigma_x^2 \end{aligned}$$

On en déduit :

$$\begin{aligned} \lambda \text{ est valeur propre de } H &\Leftrightarrow H - \lambda\, I_2 \text{ non inversible} \\ &\Leftrightarrow \det(H - \lambda\, I_2) = 0 \\ &\Leftrightarrow \lambda^2 - 2\,n\,(\sigma_x^2 + \bar{x}^2 + 1)\,\lambda + 4\,n^2\,\sigma_x^2 = 0 \\ &\Leftrightarrow \lambda \text{ est racine de } Q \end{aligned}$$

où Q est le polynôme de degré 2 défini par :

$$Q(X) \;=\; X^2 - 2\,n\,(\sigma_x^2 + \bar{x}^2 + 1)\,X + 4\,n^2\,\sigma_x^2$$

Commentaire

À ce stade de l'étude de la nature d'un point critique, le calcul de $\det(H - \lambda\, I_2)$ nous fournit toujours un polynôme de degré 2 en λ. Notons le Q. Deux cas se présentent alors :

× l'expression de Q est « simple » (c'est par exemple le cas lorsque les coefficients de Q sont numériques). Dans ce cas :

1) on détermine explicitement les racines de Q par factorisation ou calcul de discriminant.

2) les racines de Q sont les valeurs propres de H d'après les équivalences ci-dessus.

3) on en déduit le signe des valeurs propres de H et ainsi la nature du point critique étudié.

× l'expression de Q est « compliquée » (c'est par exemple le cas lorsque l'expression de Q dépend de plusieurs paramètres, comme ici). Dans ce cas, **on ne cherchera pas** à déterminer les racines de Q explicitement. On procèdera de la manière suivante :

1) on justifie l'existence de valeurs propres λ_1 et λ_2 de H (la matrice H est symétrique).

2) les valeurs propres de H sont racines de Q d'après les équivalences ci-dessus. On en déduit la factorisation de Q suivante : $Q(X) = (X - \lambda_1)\,(X - \lambda_2)$

3) on identifie les coefficients des deux expressions de Q pour en déduire des relations sur λ_1 et λ_2 (elles sont appelées *relations coefficients / racines*).

4) on détermine le signe de λ_1 et λ_2 (valeurs propres de H) grâce à ces relations, et on obtient ainsi la nature du point critique étudié.

- La matrice H est une matrice symétrique (réelle). Elle est donc diagonalisable. On note λ_1 et λ_2 ses valeurs propres **éventuellement égales**.
- D'après les équivalences précédentes, on en déduit que λ_1 et λ_2 sont racines de Q. Ainsi :

$$\begin{aligned} Q(X) &= (X-\lambda_1)(X-\lambda_2) \\ &= X^2 - (\lambda_1+\lambda_2)\,X + \lambda_1\,\lambda_2 \end{aligned}$$

D'où, par définition de Q :

$$X^2 - 2\,n\,(\sigma_x^2 + \bar{x}^2 + 1)\,X + 4\,n^2\,\sigma_x^2 \;=\; X^2 - (\lambda_1+\lambda_2)\,X + \lambda_1\,\lambda_2$$

Par identification des coefficients de ces polynômes de degré 2, on en déduit le système d'équations suivant :

$$\left\{\begin{array}{rcll} \lambda_1+\lambda_2 &=& 2\,n\,(\sigma_x^2+\bar{x}^2+1) & (*) \\[2ex] \lambda_1\,\lambda_2 &=& 4\,n^2\,\sigma_x^2 & (**) \end{array}\right.$$

× L'équation $(**)$ implique : $\lambda_1\,\lambda_2 > 0$.
On en déduit que λ_1 et λ_2 ont même signe. Le point $(\hat{a},\hat{b})$ est donc un extremum local de f sur $\mathbb{R}^2$.

× L'équation $(*)$ implique : $\lambda_1+\lambda_2 > 0$.
Or λ_1 et λ_2 ont même signe. D'où : $\lambda_1 > 0$ et $\lambda_2 > 0$.

$$\boxed{\text{On en conclut que la fonction } f \text{ admet un mimimum local en } (\hat{a},\hat{b}).}$$

□

d) Établir la formule suivante : $f(\hat{a},\hat{b}) = n\,\sigma_y^2\,\big(1 - r^2(x,y)\big)$.

Démonstration.

- Soit $(a,b)\in\mathbb{R}^2$. D'après ***1.*** :

$$\begin{aligned} & f(a,b) \\ =\ & \left(\sum_{k=1}^{n} (x_k)^2\right)\ a^2 + n\,b^2 + 2\left(\sum_{k=1}^{n} x_k\right)\ a\,b - 2\left(\sum_{k=1}^{n} x_k\,y_k\right)\ a - 2\left(\sum_{k=1}^{n} y_k\right)\ b + \sum_{k=1}^{n} (y_k)^2 \\ =\ & n\,(\sigma_x^2+\bar{x}^2)\,a^2 + n\,b^2 + 2\,n\,\bar{x}\,a\,b - 2\,n\,\big(\mathrm{cov}(x,y)+\bar{x}\,\bar{y}\big)\,a - 2\,n\,\bar{y}\,b + n\,(\sigma_y^2+\bar{y}^2) \end{aligned}$$

La deuxième égalité est obtenue grâce aux relations démontrées en ***2.a)***. Ainsi :

$$f(a,b) \;=\; n\Big((\sigma_x^2+\bar{x}^2)\,a^2 + b^2 + 2\,\bar{x}\,a\,b - 2\,\big(\mathrm{cov}(x,y)+\bar{x}\,\bar{y}\big)\,a - 2\,\bar{y}\,b + (\sigma_y^2+\bar{y}^2)\Big)$$

- D'après ***2.b)*** : $\hat{b} = \bar{y} - \bar{x}\,\hat{a}$. On obtient alors :

$$\begin{aligned} & f(\hat{a},\hat{b}) \\ =\ & n\Big((\sigma_x^2+\bar{x}^2)\,\hat{a}^2 + (\bar{y}-\bar{x}\,\hat{a})^2 + 2\,\bar{x}\,\hat{a}\,(\bar{y}-\bar{x}\,\hat{a}) - 2\,\big(\mathrm{cov}(x,y)+\bar{x}\,\bar{y}\big)\,\hat{a} \\ -\ & 2\,\bar{y}\,(\bar{y}-\bar{x}\,\hat{a}) + \sigma_y^2 + \bar{y}^2\Big) \\ =\ & n\big(\sigma_x^2\,\hat{a}^2 + \cancel{\bar{x}^2\,\hat{a}^2} + \bar{y}^2 - \cancel{2\,\bar{x}\,\bar{y}\,\hat{a}} + \cancel{\bar{x}^2\,\hat{a}^2} + \cancel{2\,\bar{x}\,\bar{y}\,\hat{a}} - \cancel{2\,\bar{x}^2\,\hat{a}^2} - 2\,\mathrm{cov}(x,y)\,\hat{a} - \cancel{2\,\bar{x}\,\bar{y}\,\hat{a}} \\ -\ & 2\,\bar{y}^2 + \cancel{2\,\bar{x}\,\bar{y}\,\hat{a}} + \sigma_y^2 + \bar{y}^2\big) \\ =\ & n\big(\sigma_x^2\,\hat{a}^2 - 2\,\mathrm{cov}(x,y)\,\hat{a} + \sigma_y^2\big) \end{aligned}$$

- Or, toujours d'après ***2.b)*** : $\hat{a} = \dfrac{\text{cov}(x,y)}{\sigma_x^2}$. Ainsi :

$$\begin{aligned} f(\hat{a},\hat{b}) &= n\left(\sigma_x^2\ \frac{\big(\text{cov}(x,y)\big)^2}{\sigma_x^4} - 2\,\text{cov}(x,y)\ \frac{\text{cov}(x,y)}{\sigma_x^2} + \sigma_y^2\right) \\ &= n\,\sigma_y^2\left(\frac{\big(\text{cov}(x,y)\big)^2}{\sigma_x^2\,\sigma_y^2} - 2\ \frac{\big(\text{cov}(x,y)\big)^2}{\sigma_x^2\,\sigma_y^2} + 1\right) \\ &= n\,\sigma_y^2\left(1 - \frac{\big(\text{cov}(x,y)\big)^2}{\sigma_x^2\,\sigma_y^2}\right) \end{aligned}$$

$$\boxed{\text{Et finalement : } f(\hat{a},\hat{b}) = n\,\sigma_y^2\Big(1 - \big(r(x,y)\big)^2\Big).}$$

□

3. a) Montrer que l'on a : $|r(x,y)| \leqslant 1$.

Démonstration.

- Tout d'abord, par définition de f :

$$f(\hat{a},\hat{b}) \ = \ \sum_{k=1}^{n} (\hat{a}\,x_k + \hat{b} - y_k)^2 \ \geqslant \ 0$$

- On en déduit, d'après la question précédente :

$$n\,\sigma_y^2\left(1 - \big(r(x,y)\big)^2\right) \geqslant 0$$

donc $1 - \big(r(x,y)\big)^2 \geqslant 0$ *(car : $n\,\sigma_y^2 > 0$)*

d'où $1 \geqslant \big(r(x,y)\big)^2$

ainsi $1 \geqslant |r(x,y)|$ *(par croissance de la fonction $\sqrt{\cdot}$ sur $[0,+\infty[$)*

$$\boxed{|r(x,y)| \leqslant 1}$$

□

b) Que peut-on dire du nuage de points lorsque $|r(x,y)| = 1$?

Démonstration.
Supposons : $|r(x,y)| = 1$.

- Alors : $\big(r(x,y)\big)^2 = 1$. Et donc :

$$f(\hat{a},\hat{b}) \ = \ n\,\sigma_y^2\left(1 - \big(r(x,y)\big)^2\right) \ = \ 0$$

- Par définition de f, on en déduit :

$$\sum_{k=1}^{n} (\hat{a}\,x_k + \hat{b} - y_k)^2 \ = \ 0$$

- Or une somme de carrés est nulle si et seulement si chacun des termes de la somme est nul. On en conclut :

$$\forall k \in [\![1, n]\!], \quad \hat{a}\, x_k + \hat{b} - y_k = 0$$

ou encore :

$$\forall x \in [\![1, n]\!], \quad y_k = \hat{a}\, x_k + \hat{b}$$

On en déduit que tous les points (x_1, y_1), ..., (x_n, y_n) sont situés sur la droite d'équation $y = \hat{a}\, x + \hat{b}$.

Commentaire

Cet exercice met en place une régression linéaire : le problème consiste à identifier une droite $y = a\,x + b$ qui ajuste bien le nuage de points (x_1, y_1), ..., (x_n, y_n).
L'erreur que l'on commet en utilisant la droite de régression pour prédire y_k à partir de x_k est : $y_k - (a\,x_k + b)$.
Une idée naturelle pour déterminer la valeur des coefficients a et b est donc de chercher la droite (donc le couple (a, b)) qui minimise la somme des carrés de ces erreurs :

$$\sum_{k=1}^{n} \big(y_k - (a\,x_k + b)\big)^2 \;=\; \sum_{k=1}^{n} (a\,x_k + b - y_k)^2$$

(on appelle d'ailleurs cette procédure « méthode des moindres carrés »)
On en déduit, après calculs, que l'unique droite rendant minimale l'erreur $\sum_{k=1}^{n} (a\,x_k + b - y_k)^2$ est la droite d'équation :

$$y \;=\; \hat{a}\,x + \hat{b} \;=\; \frac{\mathrm{cov}(x, y)}{\sigma_x^2}\,(x - \bar{x}) + \bar{y}$$

□