**CORRECTION**

**Variables aléatoires discrètes et produit scalaire**

# PROBLEME 1 : G2E 2024

***Lu dans le rapport de jury :*** « *Le soin apporté aux copies nous a semblé globalement satisfaisant, les résultats importants étant en général bien mis en valeur. Nous rappelons toutefois qu'une succession de calculs ne dispense jamais de rédiger un raisonnement cohérent, en explicitant la démarche adoptée, en introduisant les variables utilisées, en rappelant les hypothèses nécessaires et en concluant avec logique.*
*Comme chaque année, les candidats* ***qui n'ont pas suffisamment soigné la présentation de leur copie se sont vus retirer un nombre significatif de points*** »

### Partie A : Quelques exemples

***Lu dans le rapport de jury :*** « *Cette partie comportait deux questions d'algèbre. La première question, très simple, permettait aux candidats de s'approprier la définition de $\mathcal{S}_n^+(\mathbb{R})$ proposé en préambule (à condition de lire attentivement cette définition !). La seconde question demandait de diagonaliser en base orthonormale les matrices d'un sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$.*
*Nous avons constaté une bonne maîtrise de la recherche des éléments propres d'une matrice $2 \times 2$. Toutefois, nous avons regretté l'absence fréquente de l'utilisation de l'indication donnée dans l'énoncé demandant de chercher un vecteur propre sous la forme $(1, y, 1)$. Nous invitions donc les futurs candidats à toujours lire avec attention le sujet de façon à ne pas se précipiter dans des calculs fastidieux.* »

1. — La matrice $A$ n'est pas symétrique, donc n'est pas dans $\mathcal{S}_n^+(\mathbb{R})$.
   — La matrice $B$ est symétrique, et ses valeurs propres sont les racines de $X^2 - 1$, donc 1 et $-1$. Ainsi, $B$ n'est pas dans $\mathcal{S}_n^+(\mathbb{R})$.
   — La matrice $C$ est symétrique, et ses valeurs propres sont les racines de $X^2 - 2X$, donc 0 et 2. Ainsi, $C$ est dans $\mathcal{S}_n^+(\mathbb{R})$.
   — La matrice $D$ est symétrique, et ses valeurs propres sont les racines de $X^2 - 6X + 5$, donc 1 et 5. Ainsi, $D$ est dans $\mathcal{S}_n^+(\mathbb{R})$.

   ***Lu dans le rapport de jury :*** « *Cette première question a globalement été assez bien traitée par les candidats. Nous attendions naturellement que les candidats justifient leurs réponses en remarquant que $B$, $C$ et $D$ sont symétriques réelles et en donnant leurs spectres. Quelques candidats ont semblé toutefois persuadés que la symétrie d'une matrice s'observe uniquement par les coefficients diagonaux (qu'ils soient égaux ou pas aucun lien avec le caractère symétrique réel de la matrice considérée et ce ne sont pas, en général, les valeur propres !) et d'autres ont inutilement calculé le rang des matrices données. Une très large majorité des candidats ont calculé un déterminant qui les a mené à une équation du second degré, malheureusement par toujours résolue de façon correcte. Signalons enfin, qu'au moins pour les matrices $B$ et $C$, il était possible, avec un peu de pratique, de repérer directement les valeurs propres.* »

2. a) On pose

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Alors on a $\mathcal{M} = \overrightarrow{(}I_3, U)$, qui est donc bien un sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$.
Les matrices $I_3$ et $U$ ne sont pas colinéaires, donc forment une famille libre, et donc une base de $\mathcal{M}$.
On a donc $\dim \mathcal{M} = 2$.

b) On a pour $a, b, y \in \mathbb{R}$ :

$$M(a,b)\begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a+by \\ 2b+ay \\ a+by \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix}.$$

Ainsi, $\begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix}$ est vecteur propre de $M(a,b)$ si et seulement si $\lambda = a+by$ et $2b+ay = (a+by)y$.
Les valeurs $y = \sqrt{2}$ et $y = -\sqrt{2}$ conviennent.
Un vecteur $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ orthogonal à ces deux vecteurs doit vérifier

$$\begin{cases} x + \sqrt{2}y + z = 0 \\ x - \sqrt{2}y + z = 0 \end{cases}$$

Le vecteur $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ convient alors.
Ce vecteur est bien un vecteur propre de $M(a,b)$, associé à la valeur propre $a$, et orthogonal aux deux premiers.

c) Soient alors les vecteurs

$$u = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},\ v = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix},\ w = \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}.$$

La famille $(u, v, w)$ est alors une base orthonormée de vecteurs propres, et en posant $r = \sqrt{2}$ et $P$ la matrice dont les colonnes sont $u, v, w$, on a bien le résultat demandé.

d) La matrice $M(a,b)$ est toujours symétrique. Ainsi, elle appartient à $\mathcal{S}_n^+(\mathbb{R})$ si et seulement si $a \geqslant 0$, $a + br \geqslant 0$ et $a - br \geqslant 0$, ce qui est équivalent à $a \geqslant |b|r$.

***Lu dans le rapport de jury :*** *« A notre grande surprise, nous avons découvert que de nombreux candidats étaient en difficulté avec la notion d'espace vectoriel de matrices en confondant des matrices et des vecteurs de $\mathbb{R}^3$. Ensuite, de nombreux candidats ont préféré utiliser une méthode, probablement répétée en classe, reposant sur un calcul de déterminant. Si certains ont abouti ainsi aux résultats sans utiliser l'indication proposée, d'autres hélas se sont perdus dans des calculs interminables. Les candidats qui eux, avaient lu l'indication, sont* *parvenus presque toujours aux résultats en seulement quelques lignes. Signalons également que même les candidats qui ne sont pas parvenus à trouver les éléments propres de $M(a,b)$ mais qui ont proposé des résultats intéressants à la toute dernière question (par exemple en remarquant que $M(a,b) \in \S_3(\mathbb{R})$ ont vu cette réponse valorisée. »*

**Partie B : Matrice de Gram**

***Lu dans le rapport de jury :*** « *Cette partie a été globalement mal comprise. Les difficultés ont été essentiellement de deux ordres : soit une méconnaissance de la notion de produit scalaire et d'espace orthogonal, soit un manque d'aisance dans les manipulations de sommes (simples ou doubles). Nous conseillons donc aux futurs candidats de* ***travailler plus rigoureusement le chapitre « produit scalaire dans $\mathbb{R}^n$ »*** *apparu récemment au programme de BCPST et qui permet de riches développements. »*

3. a) Le sens direct est évident : les $e_i$ sont dans $E$, donc si $x$ est orthogonal à tous les vecteurs de $E$, alors il est orthogonal aux $e_i$.
   Supposons alors que $x$ est orthogonal à tous les $e_i$. Soit alors $y \in E$ ; la famille $(e_1, \dots, e_n)$ étant génératrice de $E$, on peut trouver des scalaires $\alpha_i$ tels que $y = \sum\limits_{k=1}^{n} \alpha_i e_i$.
   On a alors

$$\begin{aligned}(x|y) &= \left(x|\sum_{k=1}^{n} \alpha_i e_i\right) \\ &= \sum_{k=1}^{n} (x|e_i) \quad \text{par linéarité à droite} \\ &= 0 \quad \text{par hypothèse}\end{aligned}$$

   Ainsi, $x \in E^{\perp}$.

   b) Soit $X = (x_1, \dots, x_n) \in \mathbb{R}^n$.
   On a alors

$$\begin{aligned}X \in \ker(G) &\Leftrightarrow \forall i \in [\![1,n]\!],\ \sum_{k=1}^{n} (e_i|e_k) x_k = 0 \\ &\Leftrightarrow \forall i \in [\![1,n]\!],\ \left(e_i|\sum_{k=1}^{n} x_k e_k\right) = 0 \quad \text{par linéarité à droite} \\ &\Leftrightarrow \sum_{k=1}^{n} x_k e_k \in E^{\perp}\end{aligned}$$

   c) Supposons la famille $(e_1, \dots, e_n)$ libre. Soit alors $(x_1, \dots, x_n) \in \ker(G)$. On a alors par la question précédente $\sum\limits_{k=1}^{n} x_k e_k \in E^{\perp}$, mais on rappelle que $E$ est engendré par les vecteurs $\{e_1, \cdots, e_n\}$ donc $\sum\limits_{k=1}^{n} x_k e_k \in E \cap E^T = \{0_E\}$.

Ainsi, on a $\displaystyle\sum_{k=1}^{n} x_k e_k = 0$, puis $(x_1, \ldots, x_n) = 0$ par liberté.

Le noyau de $G$ est donc réduit à $\{0\}$, et donc la matrice $G$ est inversible.

Réciproquement, supposons $G$ inversible. Soit alors $(x_1, \ldots, x_n) \in \mathbb{R}^n$ tel que $\sum_{k=1}^{n} x_k e_k = 0$.

Alors $\sum_{k=1}^{n} x_k e_k = 0 \in E^{\perp}$, et donc $(x_1, \ldots, x_n) \in \ker(G) = \{0\}$.

La famille $(e_1, \ldots, e_n)$ est donc libre.

***Lu dans le rapport de jury :*** « *Seul le sens direct de la première partie de cette question a été véritablement comprise par les candidats. Manifestement, de nombreux candidats ont mal compris les natures des différents objets manipulés et ont tenu des raisonnements faux ou totalement dénués de sens. Dans la mesure où il y avait trois équivalences à démontrer, on attendait naturellement* ***un grand soin dans la rédaction du sens direct et du sens réciproque***. »

4. a) (Question admise dans ce sujet). On a alors

$$GX = \begin{pmatrix} \sum_{k=1}^{n} (e_1|e_k) x_k \\ \vdots \\ \sum_{k=1}^{n} (e_n|e_k) x_k \end{pmatrix},$$

puis

$$^{t}XGX = \sum_{i=1}^{n} \sum_{k=1}^{n} (e_i|e_k) x_i x_k = (x'|x')$$

où $x' = \sum_{k=1}^{n} x_k e_k$.

b) On a d'une part

$$^{t}XGX = {}^{t}X\lambda X = \lambda(X|X).$$

D'autre part, on a $^{t}XGX = (x'|x')$ d'après la question précédente.

En notant que $(X|X) > 0$ et $(x'|x') \geqslant 0$ car le produit scalaire est défini positif, on a donc $\lambda = \frac{(x'|x')}{(X|X)} \geqslant 0$.

Comme $G$ est symétrique par symétrie du produit scalaire, on a bien $G \in \mathcal{S}_n^{+}(\mathbb{R})$.

***Lu dans le rapport de jury :*** « *Le début de cette question as posé moins de difficultés : la plupart des étudiants ont su calculer le produit matriciel $GX$ et un certain nombre d'entre eux sont aussi parvenus à exprimer $X^T GX$ à l'aide d'une somme double. Seuls les meilleurs ont explicité le vecteur $x'$ afin d'en déduire $Sp(G) \subset \mathbb{R}_+$.* »

## Partie C : Matrice de covariance

***Lu dans le rapport de jury :*** « *Cette dernière partie du problème comportait deux questions : la première visait à démontrer la linéarité ) droite de la covariance et la seconde à reprendre une*

*technique abordée précédemment de façon à prouver que Sp*$(\sum) \subset \mathbb{R}_+$.
*Nous constatons que les propriétés usuelles de l'espérance et de la covariance sont maîtrisées, la toute dernière question, probablement la plus délicate du problème, n'a été que rarement comprise.* »

5. a) Soient donc $i, j, k \in [\![1, n]\!]$ et $x \in \mathbb{R}$.

$$\begin{aligned}\mathrm{Cov}(X_i, X_j + xX_k) &= \mathbb{E}(X_i(X_j + xX_k)) - \mathbb{E}(X_i)\mathbb{E}(X_j + xX_k)\\ &= \mathbb{E}(X_iX_j + xX_iX_k) - \mathbb{E}(X_i)\mathbb{E}(X_j + xX_k)\\ &= \mathbb{E}(X_iX_j) + x\mathbb{E}(X_iX_k) - \mathbb{E}(X_i)\mathbb{E}(X_j) - x\mathbb{E}(X_i)\mathbb{E}(X_k)\\ &= \mathrm{Cov}(X_i, X_j) + x\mathrm{Cov}(X_i, X_k)\end{aligned}$$

b) Soient $i \in [\![1, n]\!]$ et $x_1, \ldots, x_n \in \mathbb{R}$. Procédons par récurrence sur $n$ :
— pour $n = 1$, le résultat est donné par la question précédente.
— soit $n \in \mathbb{N}^*$ ; on suppose le résultat vrai pour ce $n$. Soient $i \in [\![1, n+1]\!]$ et $x_1, \ldots, x_{n+1} \in \mathbb{R}$. On a alors

$$\begin{aligned}\mathrm{Cov}\left(X_i, \sum_{j=1}^{n+1} x_jX_j\right) &= \mathrm{Cov}\left(X_i, \sum_{j=1}^{n} x_jX_j + x_{n+1}X_{n+1}\right)\\ &= \mathrm{Cov}\left(X_i, \sum_{j=1}^{n} x_jX_j\right) + x_{n+1}X_{n+1} \quad \text{par la question précédente}\\ &= \sum_{j=1}^{n+1} x_j\mathrm{Cov}(X_i, X_j) \quad \text{par hypothèse de récurrence}\end{aligned}$$

Par récurrence, on a bien le résultat voulu.

***Lu dans le rapport de jury :*** « *La première partie de cette question n'a soulevé que de rares difficultés, la seconde partie reposait naturellement sur la première et on attendait donc une démonstration par exemple par récurrence reposant sur des arguments précis.* »

6. a) Le produit de réels étant commutatif, il est clair que la covariance est symétrique, et donc la matrice $\Sigma$ est bien symétrique réelle.

b) Soit $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ un vecteur propre de $\Sigma$, associé à la valeur propre $\lambda$.

On a alors comme en 4, en utilisant la question précédente : $Y^T\Sigma Y = \mathrm{Cov}(y', y') = \mathbb{V}(y')$, où $y' = \sum\limits_{i=1}^{n} y_iX_i$.
D'autre part, on a $Y^T\Sigma Y = \lambda(Y|Y)$, qui est strictement positif comme en 4b.
Ainsi, on a donc $\lambda = \frac{\mathrm{Cov}(y', y')}{(Y|Y)}$, qui est positif par positivité de la variance.
Comme $\Sigma$ est symétrique, on a bien $\Sigma \in \mathcal{S}_n^+(\mathbb{R})$.

***Lu dans le rapport de jury :*** « *Il était facile de justifier que* $\sum$ *est symétrique réelle (encore fallait-il le faire avec rigueur). Il était plus délicat de justifier que* $\sum \subset \mathbb{R}_+$. *Certains candidats toutefois, sans comprendre totalement la démarche, ont pu proposer des pistes pertinentes. Comme toujours, cela était valorisé.* »

## PROBLEME 2 : AGRO B 2011

### Partie A : Fonction génératrice d'une variable à valeurs dans $\mathbf{I}_n$

***Lu dans le rapport de jury :*** *« Les correcteurs ont trouvé le sujet intéressant, bien construit, de difficultés variées et de longueur raisonnable.*
*En probabilité, là où elle serait pourtant nécessaire, l'**indépendance** est souvent « **oubliée** ». Plus généralement, quand les candidats emploient un résultat de cours, ou déjà démontré, ils citent rarement les **conditions d'application**.*
*En analyse, les notions de **convergence et d'absolue convergence** ne sont que rarement maîtrisées.*
*Les **études de fonctions simples** sont l'occasion d'erreurs pour un nombre non négligeable de candidats.*
*Les quantificateurs ne sont pas rédigés, ou mal assimilés (par exemple en A.1.a)).*
*Les parties A et B ne pouvaient être correctement traitées qu'en prenant le temps de bien lire l'énoncé, et ainsi éviter de se lancer dans de longs calculs (calculs souvent maladroitement conduits et ne répondant pas toujours à l'intitulé des questions). »*

1. a) Si $X$ est à valeurs dans $I_n$, alors $\mathbb{P}(X = k) = 0$ pour tout $k > n$.
   Dès lors $\forall t \in \mathbb{R},\ g_X(t) = \sum_{k=0}^{n} a_k t^k$.
   $\boxed{g_X \text{ est un polynôme de degré au plus } n \text{ ou encore : } g_X \in \mathbb{R}_n[X]}$.
   $g_X(1) = \sum_{k=0}^{n} a_k = \sum_{k=0}^{n} \mathbb{P}(X_k) = 1$ par définition d'une variable aléatoire à valeurs dans $I_n$.

   $$\boxed{g_X(1) = 1}$$

   Par ailleurs, $g_X$ étant un polynôme, on a : $g'_X(t) = \sum_{k=1}^{n} k a_k t^{k-1}$.
   D'où $g'_X(1) = \sum_{k=1}^{n} k\mathbb{P}(X = k) = \mathbb{E}(X)$.

   ***Lu dans le rapport de jury :*** *« Certains candidats conservent des séries et parlent de polynômes de degré infini.*
   *On voit des formulations mathématiques fausses. Par exemple, pour tout $k \in I_n$, pour tout $x$, $g_x(t) = \sum_{k=0}^{n} a_k t^k$, quelques résultats fantaisistes, par exemple :*

   $$g_X(1) = a_0 + a_1 \ \mathit{ou}\ g_X(t) = \frac{n+1}{2}(a_0 + a_1)$$

   *»*

   b) $g_X$ est un polynôme à coefficients réels. Si $g_X$ est donnée, par unicité des coefficients d'un polynôme, les coefficients $a_k$ sont déterminés de façon unique. Donc la loi de $X$ est connue.

***Lu dans le rapport de jury :*** « *Bien traitée dans* 50% *des copies.* »

2. a) Pour tout réel $t$, $E(t^{Z_1+Z_2}) = E(t^{Z_1}t^{Z_2}) = E(t^{Z_1})E(t^{Z_2})$ car les variables aléatoires $Z_1$ et $Z_2$ étant indépendantes, $t^{Z_1}$ et $t^{Z_2}$ sont indépendantes d'après le lemme de coalition. On a bien :

$$\boxed{\forall t \in \mathbb{R},\ g_{Z_1+Z_2}(t) = g_{Z_1}(t)g_{Z_2}(t)}$$

***Lu dans le rapport de jury :*** « *Beaucoup de candidats utilisent le résultat donné en début d'énoncé en justifiant son application par l'indépendance de $X_1$ et $X_2$ : peu d'entre eux pensent à en déduire l'indépendance de $t^{X_1}$ et $t^{X_2}$.*
*Quelques candidats justifient l'égalité $\mathbb{E}(t^{X_1}t^{X_2}) = \mathbb{E}(t^{X_1})\mathbb{E}(t^{X_2})$ par la* ***linéarité*** *de l'espérance* ***! !****.* »

b) Si $X$ suit une loi binomiale de paramètres $n$ et $p$, alors
$\forall t \in \mathbb{R},\ g_X(t) = \displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} t^k = (pt+q)^n$. (On a posé $q = 1-p$).
En dérivant, on obtient : $g'_X(t) = np(pt+q)^{n-1}$ donc
$\boxed{\mathbb{E}(X) = g'_X(1) = np}$ puisque $p+q=1$...

***Lu dans le rapport de jury :*** « *Résultat trouvé dans* 60% *des copies.* »

c) Si $Y$ suit aussi une loi binomiale de paramètres $n'$ et $p$, et si $X$ et $Y$ sont indépendantes, alors $\forall t \in \mathbb{R},\ g_{X+Y}(t) = (pt+q)^n(pt+q)^{n'} = (pt+q)^{n+n'}$.
On reconnaît la fonction génératrice d'une variable aléatoire suivant la loi binomiale de paramètres $n+n'$ et $p$.
La fonction génératrice caractérise une loi, donc $X+Y$ suit une loi binomiale de paramètres $n+n'$ et $p$.

***Lu dans le rapport de jury :*** « *Résultat trouvé correctement par* 50% *des candidats. Parmi les autres quelques uns se contentent de dire que ce résultat est dans le cours ; d'autres essayent de le retrouver à partir de loi, généralement en finissant par affirmer le résultat.* »

**Partie B : Fonction génératrice d'une variable à valeurs dans $N$**

1. $\forall t \in [-1,1],\ \forall n \in \mathbb{N},\ |a_n t^n| \leqslant a_n$.
La série de terme $\sum a_n$ converge, et sa somme vaut 1.
Le théorème de comparaison des séries à termes positifs nous permet d'affirmer que la série $\sum a_n t^n$ converge absolument.
Or la convergence absolue entraîne la convergence.

Donc $\boxed{g_X \text{ est défini sur } [-1,1].}$

$g_X(1) = \displaystyle\sum_{n=0}^{+\infty} a_n = 1$, par définition d'une variable aléatoire.

***Lu dans le rapport de jury :*** *« Question très rarement bien traitée (dix pour cent des candidats montrent correctement l'absolue convergence). »*

*La majoration $|a_k t^k| \leq |t^k|$ ne permet de montrer l'absolue convergence que lorsque $|t| < 1$. Le cas $|t| = 1$ doit alors être étudié à part, ce que les candidats remarquent très rarement. La solution la plus rapide consiste à effectuer la majoration $|a_k t^k| \leq a_k$. Quelques candidats se lancent dans des calculs avec les séries !!!*

*Parmi les erreurs rencontrées dans plusieurs copies, on trouve :*

— La série de terme général $a_k|t^k|$ est une série géométrique...
— $\underset{k\to+\infty}{lim} |a_k t^k| = 0$ donc la série est absolument convergente.
— $a_k|t^k|$ est majoré par 1, donc la série est absolument convergente.
— Certains écrivent : $\left|\displaystyle\sum_{k=0}^{\infty} a_k t^k\right| \leq \displaystyle\sum_{k=0}^{\infty} a_k|t^k|$ pour déduire la convergence de la série.

2. Si $X$ et $Y$ sont indépendantes, alors $\forall t \in [-1,1]$ les variables aléatoires $t^X$ et $t^Y$ sont indépendantes et admettent des espérances d'après le 1.
Donc $g_{X_1+X_2}(t) = E(t^{X_1+X_2}) = E(t^{X_1}t^{X_2}) = E(t^{X_1})E(t^{X_2}) = g_{X_1}(t)g_{X_2}(t)$.
$$\boxed{g_{X_1+X_2}(t) = g_{X_1}(t)g_{X_2}(t)}$$

3. ***Lu dans le rapport de jury :*** « ***Cette question a permis aux élèves moyens, mais sérieux, de faire la différence.*** »

a) Si $X$ suit une loi géométrique de paramètre $p$ à valeurs dans $\mathbb{N}^*$, alors
$\forall t \in [-1,1],\ g_X(t) = \displaystyle\sum_{k=1}^{+\infty} q^{k-1}pt^k = pt\sum_{k=0}^{+\infty}(qt)^k = \frac{pt}{1-qt}$, car $|qt| < 1$.

***Lu dans le rapport de jury :*** *« Attention : $X(\Omega) = \mathbb{N}^*$. Trop de candidats ont commencé la somme à 0. Trop ont oublié ou n'ont pas justifié correctement la convergence de la série géométrique. »*

b) Si $X$ suit une loi Poisson de paramètre $\lambda$, alors $\forall t \in [-1,1],\ g_X(t) = \displaystyle\sum_{k=0}^{+\infty} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}t^k = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda t}$.

***Lu dans le rapport de jury :*** *« Souvent bien traitée. »*

c) Les fonctions $g_1(t) = \dfrac{pt}{1-qt}$ et $g_2(t) = e^{\lambda(t-1)}$ sont dérivables sur $[-1,1]$, la première comme produit et inverse d'une fonction continue qui ne s'annule pas sur $[-1,1]$, la seconde par composition de fonctions continues sur cette intervalle. En dérivant en $x = 1$, on obtient :

$$g_1'(t) = \frac{p}{(1-qt)^2} \text{ soit } g_1'(1) = \frac{p}{(1-q)^2} = \frac{1}{p} = \mathbb{E}(X) \text{ si } X \hookrightarrow \mathcal{G}(p)$$

et

$$g_2'(t) = \lambda e^{\lambda(t-1)} \text{ soit } g_2'(1) = \lambda = \mathbb{E}(X) \text{ si } X \hookrightarrow \mathcal{P}(\lambda)$$

**Partie C : Fonction génératrice de la somme d'un nombre aléatoire de variables aléatoires**

1. Soit $\mathcal{P}_n$ la propriété : « $\forall t \in [0,1],\ \psi_n(t) = (f(t))^n$ ».
   - $\mathcal{P}_1$ est vraie.
   - Supposons que $\mathcal{P}_n$ soit vraie. Sous cette hypothèse, les variables aléatoires $X_1, \dots, X_{n+1}$ étant indépendantes, $X_1 + \cdots + X_n$ et $X_{n+1}$ sont indépendantes (lemme de coalition). D'après B.2 la fonction génératrice de $X_1 + \cdots + X_n + X_{n+1}$ est $\psi_{n+1} = g_{X_1+\cdots+X_n} \cdot g_{X_{n+1}}$ En utilisant l'hypothèse de récurrence, $\forall t \in [-1,1],\ \psi_{n+1}(t) = (f(t))^n f(t) = (f(t))^{n+1}$. Donc $\mathcal{P}_{n+1}$ est vraie.
   - $\boxed{\text{Par récurrence, } \forall n \in \mathbb{N}^*,\ \mathcal{P}_n \text{ est vraie.}}$

     ✐ Noter qu'une récurrence n'est pas indispensable et qu'on peut aussi dire que :
     $$\text{Si } n \in I_s^*,\ \psi_n(t) = \mathbb{E}(t^{S_n}) = \mathbb{E}\left(t^{X_1+\cdots+X_n}\right) = \mathbb{E}\left(\prod_{k=1}^{n} t^{X_k}\right)$$
     Ce qui permet de conclure par application du lemme de coalition qui assure que les $t^{X_k}$ sont indépendantes...

   - Enfin par hypothèse, $S_0 = 0$ donc $\psi_0(t) = g_{S_0}(t) = \mathbb{P}(X=0)t^0 = 1 = (f(t))^0$.

   - ***Conclusion*** : $\boxed{\forall n \in I_s,\ \psi_n(t) = (f(t))^n}$

     ***Lu dans le rapport de jury :*** « 55% *des candidats obtiennent des points sur cette question. Le cas* $n=0$ *est oublié dans la plupart des copies. Les récurrences correctes sont trop rares. Dans plusieurs copies on trouve l'erreur suivante :*
     $$\mathbb{P}\left(\sum_{i=1}^{n} X_i = n\right) = \prod_{i=1}^{n} \mathbb{P}(X_i = n)$$
     »

2. La variable aléatoire $N$ étant à valeurs dans $I_s$, la famille $\{(N=n), n \in I_s\}$ est un système complet d'événements. Utilisons la formule des probabilités totales :

   $\forall k \in \mathbb{N},\ \mathbb{P}(Y=k) = \displaystyle\sum_{n=0}^{s} \mathbb{P}((Y=k) \cap (N=n)).$

   Or si $(N=n)$ est réalisé, alors $(Y=S_n)$. Donc $\mathbb{P}(Y=k) = \displaystyle\sum_{n=0}^{s} \mathbb{P}((S_n=k) \cap (N=n)).$

   ***Lu dans le rapport de jury :*** « *Traitée convenablement par* 50% *des candidats.*
   *On trouve quelques formulations incorrectes : Si* $(N=n)$*, alors*
   $$\mathbb{P}(Y=k) = \mathbb{P}(S_n=k)$$
   »

3. $\forall t \in [-1,1],\ g(t) = g_Y(t) = \displaystyle\sum_{k=0}^{+\infty} P(Y=k)t^k = \sum_{k=0}^{+\infty} \left( \sum_{n=0}^{s} P((S_n = k) \cap (N=n)) \right) t^k$

   $g(t) = \displaystyle\sum_{k=0}^{+\infty} \left( \sum_{n=0}^{s} P(S_n = k)P(N=n) \right) t^k$

   car $N$ est indépendante des variables $X_n$, donc de $S_n$.

4. Nous sommes confrontés à une série double à termes positifs dont on sait qu'on peut, en cas de convergence, intervertir l'ordre des sommations...

   Pour tout $t \in [-1,1]$, nous avons vu en B.1. que pour tout $n \in I_s$ la série $\displaystyle\sum_{k \geq 0} P(S_n = k)t^k$ est absolument convergente donc convergente.

   Donc la combinaison linéaire $\displaystyle\sum_{n=0}^{s} P(N=n) \left( \sum_{k=0}^{\infty} P(S_n = k)t^k \right)$ est convergente et par interversion des sommes, on a :

   $\displaystyle\sum_{n=0}^{s} P(N=n) \left( \sum_{k=0}^{+\infty} P(S_n = k)t^k \right) = \sum_{k=0}^{+\infty} \left( \sum_{n=0}^{s} P(N=n)P(S_n = k)t^k \right) = g(t)$

   Donc $g(t) = \displaystyle\sum_{n=0}^{s} P(N=n) \left( \sum_{k=0}^{+\infty} P(S_n = k)t^k \right) = \sum_{n=0}^{s} P(N=n)\psi_n(t)$

   Et en utilisant C.1. on obtient :

   $$g(t) = \sum_{n=0}^{s} P(N=n)(f(t))^n = h(f(t))$$

   Ce qui montre que :

   $$\boxed{\forall \in [-1,1],\ g(t) = (h \circ f)(t).}$$

   ***Lu dans le rapport de jury :*** « *Le calcul est trouvé par* 40% *des candidats. Mais l'interversion de somme n'est quasiment jamais justifiée.* »

5. On a vu au B.3.a. que $\forall t \in [-1,1],\ f(t) = \dfrac{pt}{1-qt}$ et $h(t) = \dfrac{p't}{1-q't}$.

   Donc $g(t) = h(f(t)) = \dfrac{p'f(t)}{1-q'f(t)} = \dfrac{p'\dfrac{pt}{1-qt}}{1-q'\dfrac{pt}{1-qt}} = \dfrac{pp't}{1-(q+q'p)t}$.

   $$\boxed{g(t) = \frac{pp't}{1-(q+q'p)t}.}$$

   Or $q + q'p = 1 - p + (1-p')p = 1 - pp'$, donc $g(t) = \dfrac{pp't}{1-(1-pp')t}$. On reconnaît la fonction génératrice d'une variable aléatoire suivant une loi géométrique de paramètre $pp'$.

   $$\boxed{Y \text{ suit la loi géométrique de paramètre } pp'}$$

***Lu dans le rapport de jury :*** « *Question traitée par* 15% *des candidats, la plupart du temps de façon correcte.* ***Elle a permis à certains de faire la différence.*** »

**Partie D : Multiplication d'une bactérie**

1. Si à la génération $n$ il n'y a plus de bactéries, alors à génération $n+1$ il n'y a pas non plus de bactérie.
   L'événement $Y_n = 0$ entraîne l'événement $Y_{n+1} = 0$, donc $x_n \leqslant x_{n+1}$.
   La suite $(x_n)$ est croissante, majorée par 1 (ce sont des probabilités), donc converge.

   ***Lu dans le rapport de jury :*** « *Rarement bien traitée. Certains candidats ont essayé dès la première question de calculer $x_n$ : une lecture complète de l'énoncé aurait évité cette perte de temps.*
   *L'erreur la plus fréquente a consisté à affirmer : « le nombre de bactéries augmente au fur et à mesure des générations, donc la probabilité de disparition des bactéries diminue dans le temps ».* »

2. $Y_1$ est le nombre de fils de la bactérie de départ. Donc $Y_1$ suit une loi de Poisson de paramètre $\lambda$. Dès lors :
   $$x_1 = \mathbb{P}(Y_1 = 0) = e^{-\lambda}\frac{\lambda^0}{0!}$$
   et par définition de $f$ :
   $$f(x_0) = f(0) = e^{\lambda(x_0-1)} = e^{-\lambda}$$
   $$\boxed{x_1 = f\left(x_0\right)}$$

   ***Lu dans le rapport de jury :*** « *Traitée correctement par* 50% *des candidats.* »

3. Si $Y_1 = 0$, alors $Y_2 = 0$
   Sinon : la bactérie de départ a $Y_1$ fils qu'on peut numéroter de 1 à $Y_1$.
   Appelons $X_k$ le nombre de fils du fils numéro $k$.
   Le nombre de fils de la seconde génération est $Y_2 = X_1 + X_2 + \cdots + X_{Y_1} = S_{Y_1}$
   Par hypothèse les variables aléatoires $(X_k)_{k\geq 1}$ sont indépendantes et suivent la même loi que $X$.
   La variable $Y_1$ est aussi indépendante des variables $(X_n)_{n\geq 1}$.
   Les hypothèses sont vérifiées pour appliquer la question C.4 dans laquelle nous avons montré que :
   $$g_{Y_2} = g_{Y_1} \circ f$$
   Et comme $Y_1 \hookrightarrow \mathcal{P}(\lambda)$, on a :
   ***Conclusion*** : $\boxed{g_{Y_2} = g_{Y_1} \circ f = f \circ f.}$

   ***Lu dans le rapport de jury :*** « *Traitée par plus de* 50% *des candidats. Lorsque cette question est abordée le début est généralement bien traité.* »

4. a) Soit $\mathcal{P}_n$ la proposition : "$g_{Y_n} = f^n$" (composée $n$-ème de $f$).
   - $\mathcal{P}_1$ est vraie.
   - Supposons que $\mathcal{P}_n$ soit vraie.
     Si $Y_n = 0$, alors $Y_{n+1} = 0$
     Sinon : à la génération $n$ il y a $Y_n$ bactéries qu'on peut numéroter de 1 à $Y_n$.
     Appelons $X_k$ le nombre de fils du fils numéro $k$.
     Notons, comme dans le C. $S_n = \sum_{k=1}^{n} X_k$.
     Le nombre de fils de la $(n+1)$-ème génération est $Y_{n+1} = S_{Y_n}$
     Par hypothèse les variables aléatoires $(X_k)_{k\geq 1}$ sont indépendantes et suivent la même loi que $X$.
     Nous avons montré à la question C.4 que alors $g_{Y_{n+1}} = g_{Y_n} \circ f$
     En utilisant l'hypothèse de récurrence, $g_{Y_{n+1}} = f^n \circ f = f^{n+1}$.
   - $\boxed{\text{Par récurrence, } g_{Y_{n+1}} = f^n \text{ pour tout } n \in \mathbb{N}^*.}$

     Donc $x_n = g_{Y_n}(0) = f^n(0) = f(f^{n-1}(0)) = f(x_{n-1})$.

     $$\boxed{\forall n \geq 1, \quad x_n = f(x_{n-1})}$$

     ***Lu dans le rapport de jury :*** « *Question difficile.* 10% *des candidats obtiennent* 1/3 *des points de la question.* »

   b) Lorsque $n$ tend vers $+\infty$, $(x_n)$ tend vers $p$.
   $(x_{n-1})$ tend aussi vers $p$. La fonction $f$ étant continue en $p$, $(f(x_{n-1}))$ tend vers $f(p)$.
   Or $x_n = f(x_{n-1})$, donc par unicité de la limite $p = f(p)$.

   $$\boxed{p = f(p).}$$

   ***Lu dans le rapport de jury :*** « *La continuité est presque toujours oubliée.* »

5. Soit $\lambda \leqslant 1$.

   $\forall t \in [0,1]$, $\varphi(t) = f(t) - t = \mathrm{e}^{\lambda(t-1)} - t$, $\varphi'(t) = \lambda \mathrm{e}^{\lambda(t-1)} - 1$.
   $\forall t \in [0,1[, t-1 < 0$ donc $\mathrm{e}^{\lambda(t-1)} < 1$, donc $\lambda\mathrm{e}^{\lambda(t-1)} < 1$ et $\varphi' < 0$ sur $[0,1]$.

| $t$ | $0$ | | $1$ |
|---|---|---|---|
| $\varphi'(t)$ | | $-$ | |
| $\varphi$ | $\mathrm{e}^{-\lambda}$ | $\searrow$ | $0$ |

   $\varphi$ est strictement décroissante de $[0,1]$ sur $[0, \mathrm{e}^{-\lambda}]$.
   Le seul zéro de $\varphi$ est 1. Or les zéros de $\varphi$ sont les point fixes de $f$, donc nécessairement $p = 1$.

   $$\boxed{\text{La probabilité que la bactérie disparaisse est 1}}$$

***Lu dans le rapport de jury :*** « *Des horreurs très dommageables ! On aimerait que les étudiants prennent le temps de faire les calculs. Même quand les variations sont correctes, la décroissance stricte est très rarement citée pour justifier que* $p = 1$ »

6. Soit $\lambda > 1$.

a) $\forall t \in ]1, +\infty[, \ \theta'(t) = \dfrac{\ln u - 1}{u^2}$.

| $u$ | $1$ | | $\mathrm{e}$ | | $+\infty$ |
|---|---|---|---|---|---|
| $\theta'(t)$ | | $-$ | $0$ | $+$ | |
| $\theta$ | $1$ | $\searrow$ | $1 - \mathrm{e}^{-1}$ | $\nearrow$ | $1$ |

D'après le tableau de variations : $\forall u > 1, \ \theta(u) > 0$, donc $\dfrac{\ln u}{u} < 1$.

***Conclusion*** : $\boxed{\forall u > 1, \ \ln u < u}$
et par ailleurs

$$\ln(u) - u < 0 \Leftrightarrow e^{\ln(u) - u} < 1 \Leftrightarrow \boxed{ue^{-u} < 1}$$

***Lu dans le rapport de jury :*** « *Traitée correctement par* 40% *des candidats.* »

b) $\forall t \in [0,1], \ \varphi(t) = f(t) - t = \mathrm{e}^{\lambda(t-1)} - t$, $\varphi'(t) = \lambda \mathrm{e}^{\lambda(t-1)} - 1$ et $\varphi''(t) = \lambda^2 \mathrm{e}^{\lambda(t-1)} > 0$.
$\varphi'$ est continue, strictement croissante sur $[0,1]$ dans $J = [\lambda \mathrm{e}^{-\lambda} - 1, \lambda - 1]$ donc réalise une bijection entre ces deux intervalles.
On a montré à la question précédente que $\ln \lambda < \lambda$, donc $\lambda < \mathrm{e}^{\lambda}$ et $\lambda \mathrm{e}^{-\lambda} - 1 < 0$.
Comme $\lambda - 1 > 0$, $0$ est élément de $J$. $\boxed{\text{Il existe donc un unique } \beta \in ]0,1[ \text{ tel que } \varphi'(\beta) = 0.}$
$\varphi'$ est négative sur $[0, \beta]$ et positive sur $[\beta, 1]$.

| $t$ | $0$ | | $\alpha$ | | $\beta$ | | $1$ |
|---|---|---|---|---|---|---|---|
| $\varphi'(t)$ | $\lambda \mathrm{e}^{-\lambda} - 1 < 0$ | | $-$ | | $0$ | $+$ | |
| $\varphi$ | $\mathrm{e}^{-\lambda}$ | $\searrow$ | $0$ | $\searrow$ | | $\nearrow$ | $0$ |

***Conclusion*** : $\boxed{\varphi \text{ est strictement décroissante sur } [0, \beta] \text{ et strictement croissante sur } [\beta, 1].}$

***Lu dans le rapport de jury :*** « *Rarement abordée* »

c) $\varphi(1)$ étant égal à $0$, nécessairement $\varphi(\beta) < 0$.
La restriction de $\varphi$ à $[0, \beta]$ réalise une bijection entre $[0, \beta]$ et $[\varphi(\beta), \mathrm{e}^{-\lambda}]$. Il existe donc un réel unique $\alpha \in ]0, \beta[$ tel que $\varphi(\alpha) = 0$. Or $\varphi(\alpha) = 0$ équivaut à $f(\alpha) = \alpha$.
Donc $\boxed{\text{il existe un unique } \alpha \in ]0,1[ \text{ tel que } f(\alpha) = \alpha.}$

d) $f$ est continue strictement croissante de $[0,\alpha]$ dans $[\mathrm{e}^{-\lambda},\alpha] \subset [0,\alpha]$. Le segment $[0,\alpha]$ est stable par $f$.
Comme $x_0 = 0$, on montre facilement par récurrence que $\forall n \in \mathbb{N},\ x_n \in [0,\alpha]$.
La limite de $(x_n)$ est donc élément de $[0,\alpha]$.
Or on a vu que la limite de $(x_n)$ est un point fixe de $f$. Le seul point fixe de $f$ dans ce segment est $\alpha$, donc la suite $(x_n)$ tend vers $\alpha$.

$$\boxed{\text{La probabilité de disparition de la bactérie est } \alpha \text{ qui est strictement inférieur à 1.}}$$

7. a) *Écrivons une fonction Python* ***poisson(lbd)*** *de paramètre d'entrée le paramètre $\lambda > 0$ qui simule la réalisation d'une loi de poisson de paramètre $\lambda$ :* C'est une question de cours... il suffit de modéliser une loi binomiale de paramètres $n$ grand et $p$ petit tels que $\lambda = np$. Une rédaction possible est la suivante :

```python
def poisson(lbd):
    n = 200
    p = lbd/n
    x = 0
    for k in range(n):
        if rdm.random()<p:
            x += 1
    return x
```

b) *Écrivons une fonction* ***tempsExtinction(lbd,nBMax)*** *qui retourne la première génération qui voit disparaître la bactérie si c'est le cas et la valeur nulle si $Y$ dépasse une valeur entière* ***nBMax*** *de bactéries au-delà de laquelle on peut considérer que la population ne s'éteindra plus (par exemple* ***nBMax=100*** *semble raisonnable) :*

```python
def tempsExtinction(lbd,nBMax=100):
    Y = 1
    n = 0 # génération
    while Y != 0 and Y < nBMax:
        nB = 0
        for k in range(Y):
            nB += poisson(lbd)
        Y = nB
        n += 1
    if Y >= nBMax:
        return 0 # convention si ne s'éteint pas
    else:
        return n
```

c) *Écrivons une fonction permettant de valider les réponses fournies en 5. et 6.d) :* Il suffit de déterminer la fréquence avec laquelle la fonction précédente retourne 0, soit :

```python
def estimProbaDisparition(m,lbd):
    L = [tempsExtinction(lbd) for k in range(m)]
    return L.count(0)/m
```

**Conclusion du rapport de jury**

***Lu dans le rapport de jury :*** « *Certains élèves ont, semble-t-il, eu beaucoup de mal à entrer dans le problème et ce en raison d'une lecture trop rapide du sujet. Ils n'ont donc pas pu mettre leurs connaissances à profit.*
*Comme les années précédentes, quelques candidats rendent d'excellentes copies. On trouve aussi des copies très faibles. Ce qui donne un ensemble assez hétérogène.*
*Les correcteurs attendent clarté, concision, lecteur attentive de l'énoncé, bonne connaissance du cours (notamment ici des lois usuelles, de la notion de système complet d'événement ou d'indépendance). Ils attendent évidemment un grand soin dans les calculs élémentaires (partie D), et dans la présentation de la copie.*
*Rappelons que la malhonnêteté intellectuelle ne trompe pas le lecteur et est lourdement sanctionnée.* »