# HEC 2010

## Exercice

Soit $E = \mathbb{R}_3[X]$ l'espace vectoriel des polynômes de degré inférieur ou égal à 3 à coefficients réels. On confond polynôme de E et fonction polynomiale associée définie sur $\mathbb{R}$.
Soit $d$ l'application définie sur $E$ qui à tout polynôme $P$, associe le polynôme $d(P) = P'$, où $P'$ désigne la dérivée de $P$.

*1.* Rappeler sans démonstration la dimension de $E$ et la base canonique $\mathscr{B}$ de $E$.

*2.* Montrer que $d$ est un endomorphisme de $E$ et donner la matrice associée à $d$ dans la base $\mathscr{B}$.

*3.* Déterminer le noyau de $d$, $\mathrm{Ker}(d)$, l'image de $d$, $\mathrm{Im}(d)$, ainsi que leurs dimensions respectives.

*4.* Déterminer les valeurs propres de $d$ ainsi que les polynômes propres associés. L'endomorphisme $d$ est-il diagonalisable ?

On désigne par $\left(d^k\right)_{k \geqslant 0}$, la suite d'endomorphismes de $E$ définie par : $d^0 = \mathrm{id}$, où id représente l'endomorphisme identité et, pour tout $k$ de $\mathbb{N}$, $d^{k+1} = d^k \circ d$.
Pour tout $k$ de $\mathbb{N}$, $\mathrm{Ker}(d^k)$ désigne le noyau de $d^k$.

***5.*** ***a)*** Déterminer pour tout $k$ de $[\![1, 4]\!]$, le sous-espace $\mathrm{Ker}(d^k)$ ainsi que sa dimension.
Vérifier que pour tout $k$ de $[\![1, 4]\!]$, $d\left(\mathrm{Ker}(d^k)\right) \subset \mathrm{Ker}(d^k)$.

***b)*** Soit $P$ un polynôme de degré $r$, avec $r \in [\![0, 3]\!]$. Montrer que la famille $\left(d^k(P)\right)_{0 \leqslant k \leqslant r}$ est libre.

***6.*** Dans cette question, on cherche à déterminer les sous espaces vectoriels $F$ de $E$ tels que $d(F) \subset F$.

***a)*** On suppose que $\dim(F) = 1$.
Montrer que $F$ est un sous-espace propre de $d$. En déduire $F$.

***b)*** On suppose que $\dim(F) = 2$.
Montrer qu'il existe dans $F$ un polynôme de degré supérieur ou égal à 1. En déduire $F$.

***c)*** On suppose que $\dim(F) = 3$.
On note $\tilde{d}$ l'endomorphisme de $F$ défini par : pour tout $P$ de $F$, $\tilde{d}(P) = d(P)$.
Montrer que $\left(\tilde{d}\right)^3 = 0_{\mathscr{L}(F)}$. En déduire $F$.

## Problème

- 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 $\lambda$ un réel strictement positif. Soient $X_1$ et $X_2$ deux variables indépendantes de même loi exponentielle de paramètre $\lambda$ (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 $\mathbb{P}\big([X_1 \leqslant k]\big)$, pour tout $k$ de $X_1(\Omega)$.

***b)*** Calculer $\mathbb{E}(X_1 + X_2)$, $\mathbb{V}(X_1 + X_2)$, $\mathbb{E}(X_1 - X_2)$, $\mathbb{V}(X_1 - X_2)$.

***c)*** Établir la relation : $\mathbb{P}\big([X_1 = X_2]\big) = \dfrac{p}{1+q}$.

***2.*** ***a)*** Montrer que $Z$ suit la loi géométrique de paramètre $1 - q^2$. En déduire $\mathbb{E}(Z)$, $\mathbb{V}(Z)$ et $\mathbb{E}(T)$.

***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 : $\mathbb{P}\big([T = k]\big) = 2\ \mathbb{P}\big([X_1 = k]\big) - \mathbb{P}\big([Z = k]\big)$.

***c)*** Établir la formule : $\mathbb{V}(T) = \dfrac{q\ \left(2q^2 + q + 2\right)}{\left(1 - q^2\right)^2}$.

***3.*** ***a)*** Préciser $(T - Z)(\Omega)$.
Exprimer pour tout $j$ de $\mathbb{N}^*$, l'événement $[Z = j] \cap [Z = T]$ en fonction des événements $[X_1 = j]$ et $[X_2 = j]$. En déduire pour tout $j$ de $\mathbb{N}^*$, l'expression de $\mathbb{P}\big([Z = j] \cap [Z = T]\big)$.

***b)*** Montrer que pour tout couple $(j, l)$ de $(\mathbb{N}^*)^2$, on a : $\mathbb{P}\big([Z = j] \cap [T - Z = l]\big) = 2\ p^2\ q^{2j+l-2}$.

***c)*** Montrer que pour tout $k$ de $\mathbb{Z}$, $\mathbb{P}\big([X_1 - X_2 = k]\big) = \dfrac{pq^{|k|}}{1+q}$.
*(on distinguera trois cas : $k = 0$, $k > 0$ et $k < 0$)*

***d)*** En déduire la loi de la variable aléatoire $|X_1 - X_2|$.

***e)*** Établir à l'aide des questions précédentes que les variables $Z$ et $T - Z$ 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 ?

***b)*** Calculer en fonction de $q$, le coefficient de corrélation linéaire $\rho$ de $Z$ et $T$.

***c)*** Déterminer la loi de probabilité du couple $(Z, T)$.

***d)*** Déterminer pour tout $j$ de $\mathbb{N}^*$, la loi de probabilité conditionnelle de $T$ sachant l'événement $[Z = j]$.

***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'événement $[Z = j]$.
Calculer $\mathbb{E}(D_j)$.

## Partie III. Convergences

Dans les questions 1 à 4, $\lambda$ 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 $\lambda$.
On pose pour tout $n$ de $\mathbb{N}^*$ : $S_n = \sum\limits_{k=1}^{n} X_k$ 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 $\lambda$. Est-il sans biais ? Calculer la limite, lorsque $n$ tend vers $+\infty$, du risque quadratique associé à $\widehat{\lambda_n}$ en $\lambda$.

*3.* Dans cette question, on veut déterminer un intervalle de confiance du paramètre $\lambda$ au risque $\alpha$.
On note $\Phi$ la fonction de répartition de la loi normale centrée réduite, et $u_\alpha$ 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 $\lambda$ au risque $\alpha$. On note $\lambda_0$ 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 $\lambda$ au risque $\beta$ $(\beta \neq \alpha)$, tel que la longueur de cet intervalle soit $k$ $(k > 1)$ fois plus petite que celle obtenue avec le risque $\alpha$.

**a)** Justifier l'existence de la fonction réciproque $\Phi^{-1}$ de $\Phi$.
Quel est le domaine de définition de $\Phi^{-1}$ ?

**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 $\mathbb{N}^*$ : $T_n = \max(X_1, X_2, \ldots, X_n)$.
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 $x$, $F_{T_1}(x)$, $F_{T_2}(x)$, $\ldots$, $F_{T_n}(x)$.

***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 $\gamma$ 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 $\gamma$ est définie en langage **Scilab** par une constante `gamma`.
On rappelle que la fonction **Scilab** `rand()` permet de simuler la loi uniforme sur $]0,1[$.