Retour sur le programme (1). Petits problèmes et exercices de synthèse.

Problème 1 - Extrait de EDHEC 2008

Notions abordées

- Inégalité des accroissements finis
- Intégration sur un segment. Sommes de Riemann

$$\frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) \underset{n\to+\infty}{\longrightarrow} \int_0^1 f(t)\mathrm{d}t.$$

- Manipulation de factorielles. Récurrences.
- Théorème de transfert
- Loi conditionnelle

Partie 1. Résultats préliminaires

Soit f une fonction de classe $\mathcal{C}^1$ sur $[0;1]$.

(1) Montrer qu'il existe $M>0$ tel que, pour tous $x,y\in[0;1]$,

$$|f(x)-f(y)|\leq M|x-y|.$$

(2) En déduire que, pour tout entier $n\geq 1$, pour tout entier $k\in[\![0;n-1]\!]$, et pour tout réel $t\in\left[\frac{k}{n};\frac{k+1}{n}\right]$,

$$\left|f(t)-f\left(\frac{k}{n}\right)\right|\leq M\left(t-\frac{k}{n}\right).$$

(3) En intégrant l'inégalité précédente, montrer que, pour tout entier $n\geq 1$, pour tout entier $k\in[\![0;n-1]\!]$,

$$\left|\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\mathrm{d}t-\frac{1}{n}f\left(\frac{k}{n}\right)\right|\leq\frac{M}{2n^2}.$$

(4) En déduire que, pour tout entier $n\geq 1$,

$$\left|\int_0^1 f(t)\mathrm{d}t-\frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right)\right|\leq\frac{M}{2n}$$

puis que

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right)=\int_0^1 f(t)\mathrm{d}t.$$

(5) **Application.** Déterminer

$$\lim_{n\to+\infty} \frac{\sqrt{1}+\sqrt{2}+\cdots+\sqrt{n-1}}{n\sqrt{n}}$$

et

$$\lim_{n\to+\infty} n\left(\frac{1}{(n+1)^2}+\cdots+\frac{1}{(n+n)^2}\right).$$

(6) Pour tout couple (p,q) d'entiers naturels, on pose

$$I(p,q)=\int_0^1 x^p(1-x)^q\,\mathrm{d}x.$$

(a) Montrer que , pour tout couple $(p,q)\in\mathbb{N}\times\mathbb{N}$

$$I(p,q)=\frac{q}{p+1}I(p+1,q-1).$$

(b) En déduire que, pour tout couple $(p,q)\in\mathbb{N}\times\mathbb{N}$

$$I(p,q)=\frac{p!q!}{(p+q)!}I(p+q,0).$$

(c) Déterminer$I(p+q,0)$ et montrer finalement que

$$I(p,q)=\frac{p!q!}{(p+q+1)!}.$$

Partie 2. Étude d'une suite de variables aléatoires

On considère un paramètre entier $m\geq 2$ et une variable aléatoire $Y\hookrightarrow\mathcal{B}(m,p)$ (où $p\in[0;1]$).

(1) Rappeler les valeurs de l'espérance et de la variance de Y, puis à l'aide de la formule de Huygens-Kœnig, montrer que

$$E(Y(Y-1))=m(m-1)p^2.$$

(2) On considère maintenant une première suite de variables aléatoires (U_n) telle que, pour tout $n\geq 1$, U_n suit une loi uniforme sur $\left\{0;\frac{1}{n};...;\frac{n-1}{n}\right\}$, c'est à dire que, pour tout $l\in[\![0;n-1]\!]$,

$$P\left(U_n=\frac{k}{n}\right)=\frac{1}{n},$$

et une seconde suite de variables aléatoires (X_n) définie conditionnellement:

$$\text{Sachant que }\left(U_n=\frac{k}{n}\right),\quad X_n\hookrightarrow\mathcal{B}\left(m,\frac{k}{n}\right).$$

(a) Donner la loi de X_1.

(b) Soit $n\geq 2$. Montrer, à l'aide de la formule des probabilités totales, que, pour tout $i\in[\![0;m]\!]$,

$$P(X_n=i)=\frac{1}{n}\binom{m}{i}\sum_{k=0}^{n-1}\left(\frac{k}{n}\right)^i\left(1-\frac{k}{n}\right)^{m-i}.$$

(3) Utiliser la première question pour donner, sans calcul supplémentaire, la valeur de la somme

$$\sum_{i=0}^{m} i\binom{m}{i}\left(\frac{k}{n}\right)^i\left(1-\frac{k}{n}\right)^{m-i}.$$

(4) En déduire alors que

$$E(X_n)=\frac{m(n-1)}{2n}.$$

(5) En utilisant à nouveau la première question, donner sans calcul la valeur de la somme

$$\sum_{i=0}^{m} i(i-1)\binom{m}{i}\left(\frac{k}{n}\right)^{i}\left(1-\frac{k}{n}\right)^{m-i}.$$

(6) En déduire que

$$E(X_n(X_n-1)) = \frac{m(m-1)(n-1)(2n-1)}{6n^2}$$

puis que

$$V(X_n) = \frac{m(m+2)(n^2-1)}{12n^2}.$$

(7) (a) En utilisant les résultats obtenus aux deux premières questions de la première partie, calculer, pour tout $i \in X_n(\Omega)$,

$$\lim_{n\to+\infty} \mathrm{P}\left(X_n = i\right).$$

(b) En déduire que la suite (X_n) converge en loi vers une variable aléatoire X dont on précisera la loi.

(c) Vérifier que

$$\lim_{n\to+\infty} E\left(X_n\right) = E\left(X\right), \quad \text{et} \quad \lim_{n\to+\infty} V(X_n) = V(X).$$

Problème 2 - D'après HEC 2015

Notions abordées

- Loi géométrique. Loi exponentielle.
- Loi du max, loi du min.
- Couples de v.a. Covariance.
- Convergence en loi.
- Intervalle de confiance.
- Estimateurs. Biais.
- `(SciLab)` Simulation de loi par inversion.

Partie I. Loi exponentielle

(1) (a) Rappeler la valeur de $I_0 = \int_0^{+\infty} e^{-t}\mathrm{d}t$. Établir, pour tout $n \in \mathbb{N}^*$, la convergence de l'intégrale
$$I_n = \int_0^{+\infty} t^n e^{-t}\mathrm{d}t.$$
(b) Soit $n \in \mathbb{N}^*$. A 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 $1/\lambda$). On pose
$$Y = X_1 - X_2, \quad T = \max\left(X_1, X_2\right), \quad \text{et} \quad Z = \min\left(X_1, X_2\right).$$

(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 respectives de $V\left(X_1\right)$ et de $P\left(\left[X_1 \leq x\right]\right)$, pour tout réel x.
(b) Calculer $E\left(X_1 + X_2\right)$, $V\left(X_1 + X_2\right)$, $E\left(Y\right)$, $V\left(Y\right)$.

(4) Déterminer pour tout réel z, $F_Z\left(z\right)$ et $f_Z\left(z\right)$. Reconnaître la loi de Z, puis en déduire $E\left(Z\right)$ et $V\left(Z\right)$.

(5) (a) Montrer que pour tout réel t, on a
$$F_T(t) = \begin{cases} \left(1-e^{-\lambda}\right)^2, & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$
(b) Exprimer pour tout réel t, $f_T\left(t\right)$.
(c) Justifier l'existence de $E\left(T\right)$ et $V\left(T\right)$. Montrer, à l'aide de changements de variables affines, que
$$E\left(T\right) = \frac{3}{2\lambda}, \quad \text{et} \quad V\left(T\right) = \frac{5}{4\lambda^2}.$$

(6) On note r le coefficient de corrélation linéaire de Z et T. Montrer que $r = 1/\sqrt{5}$.

(7) (a) Préciser $Y\left(\Omega\right)$ et $|Y|\left(\Omega\right)$.
(b) Déterminer une densité de la variable aléatoire $-X_2$.
(c) Montrer que pour tout réel y, l'intégrale
$$\int_{-\infty}^{+\infty} f_{X_1}\left(t\right) f_{-X_2}\left(y-t\right)\mathrm{d}t$$
est convergente et qu'elle vaut $\frac{\lambda}{2}e^{-\lambda|y|}$. (*On distinguera deux cas* : $y \geq 0$ et $y < 0$.)
(d) Établir que la fonction $y \mapsto \frac{\lambda}{2}e^{-\lambda|y|}$ 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, $f_{|Y|}\left(y\right)$. Reconnaître la loi de $|Y| = T - Z$.

Partie II. Loi géométrique

Soient p un réel de $]0,1[$ et $q = 1 - p$. Soient X_1 et X_2 deux variables indépendantes de même loi géométrique de paramètre p (d'espérance $1/p$).

On pose :

$$Y = X_1 - X_2, \quad T = \max(X_1, X_2), \quad \text{et} \quad Z = \min(X_1, X_2).$$

On rappelle que

$$T + Z = X_1 + X_2 \quad \text{et que} \quad T - Z = |X_1 - X_2| = |Y|.$$

(8) (a) Rappeler sans démonstration les valeurs respectives de $V(X_1)$ et de $P([X_1 \leq k])$, pour tout k de $X_1(\Omega)$.
(b) Calculer $E(X_1 + X_2)$, $V({}_1 + X_2)$, $E(X_1 - X_2)$, et $V(X_1 - X_2)$.
(c) Établir la relation

$$P([X_1 = X_2]) = \frac{p}{1+q}.$$

(9) (a) Montrer que Z suit la loi géométrique de paramètre $1 - q^2$. En déduire $E(Z)$, $V(Z)$ et $E(T)$.
(b) Soit $k \in \mathbb{N}^*$. Justifier l'égalité

$$[Z = k] \cup [T = k] = [X_1 = k] \cup [X_2 = k].$$

(c) En déduire la relation suivante

$$P(T = k) = 2P(X_1 = k) - P(Z = k).$$

(d) Établir la formule :

$$V(T) = \frac{q(2q^2 + q + 2)}{(1 - q^2)^2}.$$

(10) (a) Préciser $(T - Z)(\Omega)$. Exprimer, pour tout $j \in \mathbb{N}^*$, l'évènement $[Z = j] \cap [Z = T]$ en fonction des évènements $[X_1 = j]$ et $[X_2 = j]$. En déduire, pour tout $j \in \mathbb{N}^*$, l'expression de $P([Z = j] \cap [Z = T])$.
(b) Montrer que pour tout couple $(j, l) \in (\mathbb{N}^*)^2$, on a

$$P([Z = j] \cap [T - Z = l]) = 2p^2 q^{2j+l-2}.$$

(c) Montrer, en distinguant les trois cas $k = 0$, $k > 0$ et $k < 0$, que, pour tout $k \in mathbbZ$,

$$P([X_1 - X_2 = k]) = \frac{pq^{|k|}}{1+q}.$$

(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.

(11) (a) A l'aide du résultat de la Question 10.e, calculer $\operatorname{cov}(Z, T)$. Les variables Z et T sont-elles indépendantes ?
(b) Calculer en fonction de q, le coefficient de corrélation linéaire ρ de Z et T.
(c) Déterminer la loi de probabilité du couple (Z, T).
(d) Déterminer pour tout $j \in \mathbb{N}^*$, la loi de probabilité conditionnelle de T sachant l'évènement $[Z = j]$.
(e) Soit $j \in \mathbb{N}^*$. 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 $E(D_j)$.

Partie III. Convergences

Dans les questions 12 à 15, λ désigne un paramètre réel strictement positif, **inconnu**.

Pour $n \in \mathbb{N}^*$, on considère un n-échantillon $(X_1, X_2, ..., X_n)$ de variables aléatoires à valeurs strictement positives, indépendantes, de même loi exponentielle de paramètre λ.

On pose, pour tout $n \in \mathbb{N}^*$,

$$S_n = \sum_{k=1}^{n} X_k, \quad \text{et} \quad J_n = \lambda S_n.$$

(12) Calculer, pour tout $n \in \mathbb{N}^*$, $E(S_n)$, $V(S_n)$, $E(J_n)$ et $V(J_n)$.

(13) On admet qu'une densité f_{J_n} de J_n est donnée par la formule

$$f_{J_n}(x) = \begin{cases} \frac{e^{-x}x^{n-1}}{(n-1)!}, & \text{si } x > 0 \\ 0, & \text{si } x \leq 0 \end{cases}.$$

(a) À l'aide du théorème de transfert, établir pour tout entier $n \geq 3$, l'existence de $E\left(\frac{1}{J_n}\right)$ et de $E\left(\frac{1}{J_n^2}\right)$, et donner leur valeurs respectives.

(b) On pose, pour tout entier $n \geq 3$,

$$\widehat{\lambda_n} = \frac{n}{S_n}.$$

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.

(14) 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

$$\Phi(u_\alpha) = 1 - \frac{\alpha}{2}.$$

(a) Enoncer le théorème de la limite centrée. En déduire que la variable aléatoire N_n définie par $N_n = \lambda \frac{S_n}{\sqrt{n}} - \sqrt{n}$ converge en loi vers la loi normale centrée réduite.

(b) En déduire que pour n assez grand, on a approximativement

$$P([-u_\alpha \leq N_n \leq u_\alpha]) = 1 - \alpha.$$

(c) Montrer que pour n assez grand, l'intervalle

$$\left[\left(1 - \frac{u_\alpha}{\sqrt{n}}\right)\widehat{\lambda_n}, \left(1 + \frac{u_\alpha}{\sqrt{n}}\right)\widehat{\lambda_n}\right]$$

est un intervalle de confiance de λ au risque α. On note λ_0 la réalisation de $\widehat{\lambda_n}$ sur le n-échantillon.

(15) Avec le n-échantillon $(X_1, X_2, ..., X_n)$, 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

$$\beta = 2\Phi\left(\frac{1}{k}\Phi^{-1}(\alpha/2)\right).$$

(c) En déduire que $\beta > \alpha$. Ce dernier résultat était-il prévisible?

Dans les questions 16 à 18, on suppose que $\lambda = 1$.

(16) On pose, pour tout $n \in \mathbb{N}^*$,
$$T_n = \max\left(X_1, X_2, ..., X_n\right).$$
Pour tout $n \in \mathbb{N}^*$, pour tout réel $x \geq 0$, on pose
$$g_n(x) = \int_0^x F_{T_n}(t)\,\mathrm{d}t \quad \text{et} \quad h_n(x) = \int_0^x t f_{T_n}(t)\,\mathrm{d}t.$$
(a) Exprimer $h_n(x)$ en fonction de $F_n(x)$ et $g_n(x)$.
(b) Déterminer, pour tout réel t, l'expression de $F_{T_n}(t)$ en fonction de t.
Établir, pour tout entier $n \geq 2$, la relation :
$$g_{n-1}(x) - g_n(x) = \frac{1}{n} F_{T_n}(x).$$
(c) En déduire, pour tout $n \in \mathbb{N}^*$, pour tout réel $x \geq 0$, l'expression de $g_n(x)$ en fonction de x, et de $F_{T_1}(x), F_{T_2}(x), ..., F_{T_n}(x)$.
(d) Montrer que $F_{T_n}(x) - 1$ est équivalent à $-ne^{-x}$, lorsque x tend vers $+\infty$.
(e) Déduire des questions c) et d) l'existence de $E(T_n)$ et montrer que
$$E(T_n) = \sum_{k=1}^{n} \frac{1}{k}.$$

(17) On veut étudier dans cette question la convergence en loi de la suite de variables aléatoires (G_n) dé finie, pour tout $n \in \mathbb{N}^*$, par
$$G_n = T_n - E(T_n).$$
On pose, pour tout $n \in \mathbb{N}^*$, $\gamma_n = -\ln n + E(T_n)$ et on admet *sans démonstration* que la suite (γ_n) est convergente; on note γ sa limite.

(a) Montrer que pour tout x réel et n assez grand, on a
$$F_{G_n}(x) = \left(1 - \frac{1}{n} e^{-(x+\gamma_n)}\right)^n.$$
(b) En déduire que, pour tout x réel, on a :
$$\lim_{n \to +\infty} F_{G_n}(x) = e^{-e^{-(x+\gamma)}}.$$
(c) Montrer que la fonction $F_G : \mathbb{R} \to \mathbb{R}$ définie par
$$F_G(x) = e^{-e^{-(x+\gamma)}}$$
est la fonction de répartition d'une variable aléatoire G à densité. Conclure.

(18) (a) Soit X une variable aléatoire à densité de fonction de répartition F_X strictement croissante. Déterminer la loi de la variable alétoire Y définie par $Y = F_X(X)$.
(b) Écrire une fonction, sous `SciLab`, d'en-tête `y=Gumbel()` qui permet de simuler la variable aléatoire G. On supposera que la constante γ est déjà définie dans `SciLab` et stockée sous `gamma`. On utilisera la commande `rand()` qui permet de simuler la loi uniforme sur $]0;1[$.

Exercice 1

Notions abordées

- Continuité. Dérivabilité.
- Développements limités.
- Suite d'intégrales impropres.
- Fonction de répartition de $\mathcal{N}(0;1)$.
- Changement de variable.
- `(SciLab)`. Calcul du premier entier vérifiant une condition donnée.

Enoncé

(1) Soit N la fonction définie sur l'intervalle $[0,1[$, à valeurs réelles, telle que :

$$N(x) = x^2 - 2x - 2(1-x)\ln(1-x).$$

(a) Montrer que la fonction N est de classe $\mathcal{C}^1$ sur $[0,1[$.
(b) Montrer que pour tout $x \in [0,1[$, on a : $\ln(1-x) \leq -x$.
(c) Montrer que pour tout $x \in [0,1[$, on a : $N'(x) \leq 0$.
(d) En déduire le signe de N sur l'intervalle $[0,1[$.

(2) Soit f la fonction définie sur l'intervalle $[0,1[$ par

$$f(x) = \begin{cases} -2\dfrac{x+\ln(1-x)}{x^2}, & \text{si } 0 < x < 1 \\ 1, & \text{si } x = 0 \end{cases}$$

(a) Déterminer le développement limité en 0 à l'ordre 2 de $\ln(1-x)$.
(b) En déduire $\lim_{x\to 0} f(x)$.
(c) En déduire que la fonction f est continue sur $[0,1[$.
(d) Montrer que f est dérivable sur $]0,1[$ et que pour tout $x \in]0,1[$, on a : $f'(x) = -2\dfrac{N(x)}{x^3(1-x)}$.
(e) Dresser le tableau de variation de la fonction f sur $[0,1[$.
En déduire que f réalise une bijection strictement croissante de $[0,1[$ dans $[1,+\infty[$.

(3) On pose pour tout $n \in \mathbb{N}^*$ et pour tout $x \in [0,1[$:

$$g_n(x) = \exp\left(-\frac{nx^2}{2}f(x)\right).$$

(a) Montrer, à l'aide de la question 2e, que pour tout $x \in [0,1[$, on a :

$$0 \leq g_n(x) \leq \exp\left(-\frac{nx^2}{2}\right).$$

(b) En déduire la convergence de l'intégrale $\int_0^1 g_n(x)\,\mathrm{d}x$. On pose alors, pour tout $n \in \mathbb{N}^*$,

$$I_n = \int_0^1 g_n(x)\,\mathrm{d}x.$$

(c) Montrer que pour tout $x \in [0,1[$,

$$\frac{g_{n+1}(x)}{g_n(x)} \leq 1.$$

(d) En déduire que la suite $(I_n)_{n\in\mathbb{N}^*}$ est décroissante.
(e) On note φ une densité et Φ la fonction de répartition de la loi normale centrée réduite.
Rappeler la formule définissant φ puis dresser le tableau de variation complet de Φ sur $\mathbb{R}$ en précisant la valeur de $\Phi(0)$.

(f) Montrer l'encadrement :

$$0 \leq I_n \leq \sqrt{\frac{2\pi}{n}} \left(\Phi\left(\sqrt{n}\right) - \frac{1}{2} \right).$$

On pourra utiliser la question 3a puis effectuer le changement de variable $u = x\sqrt{n}$.

(g) Montrer que pour tout $n \in \mathbb{N}^*$, on a :

$$0 \leq I_n \leq \sqrt{\frac{\pi}{2n}}.$$

(h) Écrire un programme en `SciLab` qui détermine une valeur de n pour laquelle $I_n \leq 10^{-1}$.

(i) Déterminer à présent par le calcul une valeur de n pour laquelle $I_n \leq 10^{-1}$. *On donne* $50\pi \approx 157.08$.

Exercice 2 - D'après HEC 2003

Notions abordées

- Valeurs propres.
- Matrices aléatoires.
- Covariance.

Enoncé

(1) Soit a et b deux réels strictement positifs et A la matrice carré d'ordre 2 définie par : $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$.

(a) Montrer que si a et b sont égaux, la matrice A n'est pas inversible.

(b) Calculer la matrice $A^2 - 2aA$. En déduire que, si a et b sont distincts, la matrice A est inversible et donner la matrice A^{-1}.

(c) Montrer que les valeurs propres de A sont $a+b$ et $a-b$.

(d) On pose

$$\Delta = \begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix}.$$

Déterminer une matrice Q, carrée d'ordre 2 à coefficients réels, inversible et dont les éléments de la première ligne sont égaux à 1, vérifiant $A = Q\Delta Q^{-1}$.

(e) Calculer la matrice Q^{-1} et, à l'aide de la question précédente, calculer la matrice A^n pour tout entier naturel non nul n.

(2) Soit p un réel vérifiant $0 < p < 1$ et q le réel $1-p$. On suppose que X et Y sont deux variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, indépendantes et suivant la même loi géométrique de paramètre p.

Pour tout ω de Ω, on désigne par $M(\omega)$ la matrice carrée d'ordre 2

$$M(\omega) = \begin{pmatrix} X(\omega) & Y(\omega) \\ Y(\omega) & X(\omega) \end{pmatrix}$$

et on note $S(\omega)$ (respectivement $D(\omega)$) la plus grande (respectivement la plus petite) valeur propre de $M(\omega)$ et on définit ainsi deux variables aléatoires sur $(\Omega, \mathcal{A}, P)$.

(a) Montrer que la probabilité de l'événement $[X = Y]$ est donnée par: $P([X = Y]) = \dfrac{p}{2-p}$ et en déduire la probabilité de l'événement $\{\omega \in \Omega \ ; \ M(\omega) \text{ est inversible}\}$.

(b) Calculer la covariance des variables aléatoires S et D.

(c) Calculer les probabilités $P([S = 2] \cap [D = 0])$, $\mathbf{P}([S = 2])$ et $P([D = 0])$.
Les variables aléatoires S et D sont-elles indépendantes ?

(d) Établir, pour tout entier naturel n supérieur ou égal à 2

$$P([S = n]) = (n-1)p^2q^{n-2}.$$

(e) En déduire, lorsque p est égal à $\frac{2}{21}$, que la valeur la plus probable de la plus grande valeur propre des matrices $M(\omega)$ possibles est 11.

Problème 3 - D'après HEC 2007

Notions abordées

- Espace vectoriel de polynômes
- Coordonnées dans une base

Notations

On **rappelle** les résultats suivants :

- Toute famille de polynôme de degrés échelonnés est libre.
- Tout polynôme admettant une infinité de racines est le polynôme nul.

Soit $n \in \mathbb{N}^*$. On note Δ_n l'application qui à tout polynôme P de $\mathbb{R}_n[X]$ associe le polynôme

$$\Delta_n(P) = P(X+1) - P(X).$$

On pose

$$P_0 = 1, \quad P_1 = X, \quad P_2 = \frac{1}{2}X(X-1) \quad \text{et} \quad \forall k \in [\![3;\, n]\!], \quad P_k = \frac{1}{k!}X(X-1)\dots(X-k+1).$$

Partie I : Étude de Δ_n.

(1) Justifier que Δ_n est un endomorphisme de $\mathbb{R}_n[X]$.

(2) (a) Montrer que $(P_0, \dots, P_n)$ est une base de $\mathbb{R}_n[X]$.

(b) Vérifier que :

$$\Delta_n(P_0) = 0 \quad \text{et} \quad \forall k \in [\![1;\, n]\!], \quad \Delta_n(P_k) = P_{k-1}.$$

(c) En déduire que $\operatorname{Im}(\Delta_n) = \mathbb{R}_{n-1}[X]$ puis préciser $\operatorname{rg}(\Delta_n)$.

(d) En déduire que $\operatorname{Ker}(\Delta_n)$ est l'ensemble $\mathbb{R}_0[X]$ des polynômes constants.

(e) L'application Δ_n est-elle un automorphisme de $\mathbb{R}_n[X]$?

(3) Dans le cas particulier où $n = 3$, écrire la matrice M de Δ_3 relativement à la base (P_0, P_1, P_2, P_3).

Partie II : Coordonnées d'un polynôme dans la base $(P_0, \dots, P_n)$.

Pour tout $i \in \mathbb{N}^*$, on rappelle que

$$\Delta_n^i = \underbrace{\Delta_n \circ \dots \circ \Delta_n}_{i \text{ fois}}.$$

(1) Soit $k \in [\![1;n]\!]$ et $i \in \mathbb{N}^*$. Montrer que

$$\Delta_n^i(P_k) = \begin{cases} P_{k-i} & \text{si } i \leq k, \\ 0 & \text{si } i > k. \end{cases}.$$

(2) Soit $P \in \mathbb{R}_n[X]$.

(a) Justifier qu'il existe des réels $(a_0, a_1, \dots, a_n)$ tels que $P(X) = \sum_{k=0}^n a_k P_k(X)$.

(b) En déduire que pour tout $i \in [\![1;n]\!]$

$$\Delta_n^i(P)(X) = \sum_{k=0}^{n-i} a_{k+i} P_k(X).$$

(c) En évaluant l'expression précédente en 0, montrer alors que

$$P(X) = \sum_{i=0}^{n} \Delta_n^i(P)(0)\, P_i(X).$$

Partie III : Généralisation à $\mathbb{R}[X]$.

On note Δ l'endomorphisme qui à tout polynôme P de $\mathbb{R}[X]$ associe le polynôme $\Delta(P) = P(X+1) - P(X)$.

(1) Soit $P \in \operatorname{Ker}(\Delta)$.
   (a) Montrer que pour tout $k \in \mathbb{N}$, $P(k) = P(0)$.
   (b) En considérant le polynôme, $Q(X) = P(X) - P(0)$, en déduire que $\operatorname{Ker}(\Delta)$ est l'ensemble $\mathbb{R}_0[X]$ des polynômes constants.

(2) Montrer que Δ est surjectif. *On pourra utiliser la question 2c de la partie I*

Problème 4*** - Inspiré de HEC 2004 S et Annales ESCP

Notions abordées

- Lois normales
- Lois discrètes à valeurs dans $\mathbb{Z}$
- Changement de variables.
- IAF
- Convexité
- Inégalité de Markov
- Diagonalisation

Notations

L'objet de ce problème est l'étude de diverses propriétés des lois normales. Toutes les variables utilisées sont définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

On dit qu?une variable aléatoire à valeurs dans $\mathbb{Z}$ admet une espérance si et seulement si **les** séries de terme général $nP(X = n)$ et $nP(X = -n)$ convergent et dans ce cas

$$E(X) = \lim_{n\to+\infty} \sum_{k=1}^{n} kP(X = k) - \lim_{n\to+\infty} \sum_{k=1}^{n} kP(X = -k).$$

On admet que les formules de calcul de covariance d'un couple de variables à densité sont identiques à celles d'un couple de variables discrètes.

On considère que la variable nulle suit une loi normale. On dit qu'une variable est **gaussienne** si elle suit une loi normale.

On dit qu'un vecteur $(X_1, X_2, ..., X_n)$ de variables aléatoires est **gaussien** si pour tout $(a_1, a_2, ..., a_n)$ réels, la variable $a_1X_1 + a_2X_2 + ... + a_nX_n$ suit une loi normale.

On pourra utiliser sans la démontrer la propriété suivante :
Si X et Y sont deux variables à densité indépendantes et de densités respectives f et g bornées alors $X + Y$ est une variable à densité, dont une densité h est donnée par :

$$\forall x \in \mathbb{R},\ h(x) = \int_{-\infty}^{+\infty} f(t)g(x - t)\mathrm{d}t.$$

Partie I - Stabilité par la somme

On rappelle l'expression de la densité d'une variable gaussienne de moyenne m et de variance σ^2

$$\forall x \in \mathbb{R}, \qquad f_{m,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

(1) Montrer que $f_{0,\sigma}$ est une fonction bornée sur $\mathbb{R}$.

(2) Soit a un réel strictement positif et b et c deux réels quelconques. Trouver trois réels α, m et σ, en fonction de a, b et c tels que :

$$\forall x \in \mathbb{R}, \qquad \frac{(x-m)^2}{2\sigma^2} + \alpha = ax^2 + bx + c.$$

(3) En déduire que

$$\int_{-\infty}^{+\infty} \exp(-(ax^2 + bx + c))\mathrm{d}x = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2 - 4ac}{4a}\right).$$

(4) Démontrer que $\forall a>0$, $\forall b\in\mathbb{R}$, les assertions suivantes sont équivalentes :
- X suit la loi normale de paramètres m,σ^2 ;
- $aX+b$ suit la loi normale de paramètres $am+b, a^2\sigma^2$.

(5) Soient G et $G?$ deux variables aléatoires gaussiennes, indépendantes, centrées réduites. Montrer que $G+G?$ est une variable à densité dont on déterminera une densité h. En déduire que $G+G?$ est une variable gaussienne dont on donnera l'espérance et la variance.

(6) Soient X_1 et X_2 deux variables aléatoires gaussiennes indépendantes, d'espérances respectives m_1 et m_2 et de variances respectives σ_1^2 et σ_2^2. Déduire de la question précédente que X_1+X_2 suit une loi normale dont on précisera les paramètres.

Partie II - Partie entière et partie décimale d'une variable gaussienne

Soit T une variable gaussienne centrée réduite, on pose $X=\lfloor R\rfloor$ et $Y=T-X$.

(6) Préciser $X(\Omega)$ et déterminer la loi de X en fonction de la fonction de répartition Φ de la loi normale centrée réduite.
(7) Montrer que $\forall k\in\mathbb{Z}$, $P(X=-k)=P(X=k-1)$.
(8) Montrer à l?aide de l?inégalité des accroissements finis que :

$$\forall k\in\mathbb{N},\qquad \Phi(k+1)-\Phi(k)\leq\frac{1}{\sqrt{2\pi}}e^{-k/2}.$$

(9) En déduire que $E(X)$ existe.
(10) Montrer que $\forall n\in\mathbb{N}^*$,

$$\sum_{k=1}^{n}k(P(X=k)-P(X=-k))=n(\Phi(n+1)-\Phi(n))+\Phi(0)-\Phi(n).$$

En déduire la valeur de $E(X)$.
(11) Préciser $Y()$ et déterminer $E(Y)$.
(12) Montrer que la fonction de répartition F_y de Y vérifie la relation :

$$\forall y\in[0,1[,\qquad F_Y(y)=\Phi(y)-\frac{1}{2}+\sum_{k=1}^{+\infty}k(\Phi(k+y)-\Phi(k-y)).\qquad(1)$$

(13) Déterminer $F_Y\left(\frac{1}{2}\right)$.
(14) Comparer $P([X=0]\cap[Y\leq\frac{1}{2}]$ et $P(X=0)P(Y\leq\frac{1}{2})$.
(On donne $\Phi(\frac{1}{2})\simeq 0.6915$ et $\Phi(1)\simeq 0.8413$.)
Que peut-on en déduire?

Partie III - Variables sous-gaussiennes

Pour $\alpha>0$, on dit qu'une variable X est α-sous-gaussienne si pour tout t réel,

$$E(e^{tX})\leq e^{\alpha^2t^2/2}.$$

(15) À l?aide d?un changement de variable, calculer pour tout réel t,

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{tx-x^2/2}\mathrm{d}x.$$

En déduire que si X est une variable gaussienne centrée réduite, X est 1-sous-gaussienne.
(16) Montrer que $\forall p\geq 1$, $2^pp!\leq(2p)!$

(17) En déduire que :

$$\forall t \in \mathbb{R}, \qquad \sum_{n=0}^{N} \frac{t^n}{n!} + \sum_{n=0}^{N} \frac{(-t)^n}{n!} \leq 2 \sum_{n=0}^{\lfloor N/2 \rfloor} \frac{t^{2n}}{2^n n!}$$

puis que

$$\frac{e^t + e^{-t}}{2} \leq e^{t^2/2}.$$

(18) Soit t un réel et $x \in [-1, 1]$.
En utilisant la convexité de la fonction exponentielle sur l?intervalle d?extrémités t et $-t$, montrer que

$$e^{tx} \leq \frac{1+x}{2} e^t + \frac{1-x}{2} e^{-t}.$$

(19) Soient X_1, X_2,...,X_n des variables indépendantes, α-sous-gaussiennes et $\mu_1, \mu_2, ..., \mu_n$ des réels tels que

$$\sum_{i=1}^{n} \mu_i^2 = 1.$$

Montrer que $\sum_{i=1}^{n} \mu_i X_i$ est α-sous-gaussienne.

(20) Soit X une variable α-sous-gaussienne et $\lambda > 0$.

(a) En utilisant l'inégalité de Markov, montrer que $\forall t > 0$, on a :

$$P(X \geq \lambda) \leq \exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right).$$

(b) En déduire que

$$P(|X| \geq \lambda) \leq 2\exp\left(\frac{\alpha^2 t^2}{2} - t\lambda\right).$$

(c) Étudier la fonction $t \mapsto \frac{\alpha^2 t^2}{2} - t\lambda$ et montrer que $\forall \lambda > 0$,

$$P(|X| \geq \lambda) \leq 2\exp\left(-\frac{\lambda^2}{2\alpha^2}\right).$$

Partie IV - Vecteurs gaussiens

(21) Soit $(X_1, X_2, ..., X_n)$ un vecteur gaussien. Montrer que chaque variable X_i suit une loi normale.
(22) À l'aide de la **Partie I**, montrer par récurrence que si X_1, X_2,...,X_n sont des variables aléatoires indépendantes suivant des lois normales, alors le vecteur $(X_1, X_2, ..., X_n)$ est gaussien.

(23) Dans cette question X suit la loi normale centrée réduite, et Y, indépendante de X, suit la loi discrète définie par $P(Y = 1) = P(Y = -1) = \frac{1}{2}$.
(a) Déterminer la loi de XY.
(b) Le vecteur (X, XY) est-il gaussien? On pourra calculer $P(X + XY = 0)$.

Soit $X_1, X_2, ..., X_n$ n variables aléatoires admettant chacune une variance, on appelle **matrice de variance/covariance** et on note $M(X_1, X_2, ..., X_n)$ la matrice de $M_n(\mathbb{R})$ dont les coefficients $a_{i,j}$ sont donnés par

$$\forall (i,j) \in [\![1; n]\!]^2, \qquad a_{i,j} = \mathrm{cov}(X_i, X_j).$$

On appelle **vecteur espérance** et on note $E\left(\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}\right)$ la matrice colonne de $M_{n,1}(\mathbb{R})$ dont les coefficients a_i sont donnés par

$$\forall i \in [\![1; n]\!], \qquad a_i = E(X_i).$$

(24) Justifier que $M(X_1, X_2, ..., X_n)$ est diagonalisable.

(25) Dans cette question (X, Y, Z) est un vecteur gaussien tel que

$$M(X,Y,Z) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 6 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \qquad \text{et} \qquad E\left(\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

(a) Déterminer la loi de $U = X + Y$ et celle de $V = X + Z$.

(b) Écrire la matrice de variance/covariance $M(U, V)$.

(c) Déterminer une matrice P inversible de $M_2(\mathbb{R})$ et une matrice D diagonale de $M_2(\mathbb{R})$ telle que $M(U, V) = PDP^{-1}$.

(d) Montrer que les variables U? et V? définies par

$$\begin{pmatrix} U' \\ V' \end{pmatrix} = P^{-1} \begin{pmatrix} U \\ V \end{pmatrix}$$

sont non corrélées.