
Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2014

Épreuve de mathématiques

Durée : 4h

Objectif du problème et remarques liminaires :

- On se propose ici d'introduire les notions d'applications hölderiennes et « à variation bornée ».
- Les cinq parties du problème sont relativement indépendantes les unes des autres.
- L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours.

Notations :

- Si A et B sont deux ensembles, B^A est l'ensemble des applications de A dans B .
- Dans la suite, on note E le $\mathbb{C}$ -espace vectoriel $\mathbb{C}^{\mathbb{Z}}$ muni de ses lois usuelles.
- Dans tout le problème, $T \in \mathbb{R}^{+*}$, $\omega = \frac{2\pi}{T}$ et on convient que $0^x = 0$ si $x \in \mathbb{R}^{+*}$.
- Si $(p, q) \in \mathbb{N}^2$, alors $\llbracket p, q \rrbracket$ désigne l'ensemble $\{k \in \mathbb{N}, p \leq k \leq q\}$.
- Si $(a, b) \in \mathbb{R}^2$, $a \leq b$ et $n \in \mathbb{N}^*$, un élément $x = (x_k)_{k \in \llbracket 0, n \rrbracket}$ de $[a, b]^{\llbracket 0, n \rrbracket}$ est dit subdivision de $[a, b]$ si, et seulement si, x est croissant et tel que $x_0 = a$ et $x_n = b$. Si c'est le cas et $f \in \mathbb{C}^{[a, b]}$, on note $V_x(f)$ la somme (appelée variation de f relativement à x) :

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$$

En outre, $S_{[a, b]}$ désignera l'ensemble des subdivisions de $[a, b]$. Enfin, on dit que f est à variation bornée (sur $[a, b]$) si, et seulement si, l'ensemble $\{V_x(f), x \in S_{[a, b]}\}$ est majoré et, lorsque c'est le cas, on appelle variation de f (sur $[a, b]$) sa borne supérieure, que l'on note $V_{[a, b]}(f)$.

- Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$ et $p \in \mathbb{N}$, $C_T^p(\mathbb{R}, \mathbb{K})$ est classiquement le $\mathbb{K}$ -espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{K}$ de classe C^p et T -périodiques. De plus, si $\alpha \in \mathbb{R}^{+*}$ et I est un intervalle non vide de $\mathbb{R}$, on note $H^\alpha(I, \mathbb{K})$ l'ensemble des applications f de I dans $\mathbb{K}$ qui vérifient : $\exists c \in \mathbb{R}^+ \quad \forall (x, y) \in I^2 \quad |f(y) - f(x)| \leq c \cdot |y - x|^\alpha$

On note aussi $H_T^\alpha(\mathbb{R}, \mathbb{K})$ l'ensemble des éléments de $H^\alpha(\mathbb{R}, \mathbb{K})$ T -périodiques.

- Dans tout le problème, F est l'ensemble des éléments $(z_n)_{n \in \mathbb{Z}}$ de E satisfaisant la condition :

pour tout élément p de $\mathbb{N}$, z_n est négligeable devant $|n|^{-p}$ lorsque n tend vers $-\infty$ et $+\infty$.

- Enfin, si $f \in C_T^0(\mathbb{R}, \mathbb{C})$ et $n \in \mathbb{Z}$, on note $c_n(f)$ le nombre complexe :

$$\frac{1}{T} \cdot \int_0^T (f(t) \cdot e^{-i \cdot n \cdot \omega \cdot t}) dt$$

PARTIE I

- 1/ Vérifier que F est un sous-espace vectoriel de E .
- 2/ Constater que la suite $(z^{|n|}/|n|!)_{n \in \mathbb{Z}}$ appartient à F si $z \in \mathbb{C}$.
- 3/ Soient P un élément non nul de $\mathbb{C}[X]$ et $z \in \mathbb{C}$.

a) Montrer qu'il existe un nombre entier naturel n_0 tel que :

$$\forall n \in \mathbb{Z} \quad |n| \geq n_0 \Rightarrow P(n) \neq 0.$$

b) En déduire que $(P(n).z^{|n|})_{n \in \mathbb{Z}}$ appartient à F si, et seulement si : $|z| < 1$.

4/ a) Soient $n \in \mathbb{Z}$ et $f \in C_T^1(\mathbb{R}, \mathbb{C})$. Montrer que $f' \in C_T^0(\mathbb{R}, \mathbb{C})$ et que $c_n(f') = i.n.\omega.c_n(f)$.

b) Prouver que $(c_n(f))_{n \in \mathbb{Z}} \in F$ si $f \in C_T^\infty(\mathbb{R}, \mathbb{C})$.

PARTIE II

Dans toute la suite du problème, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

5/ On suppose ici que $\alpha \in]1, +\infty[$.

a) Vérifier que, si $f \in H^\alpha(\mathbb{R}, \mathbb{K})$, alors f est dérivable et préciser f' .

b) En déduire $H^\alpha(\mathbb{R}, \mathbb{K})$ puis $H_T^\alpha(\mathbb{R}, \mathbb{K})$.

6/ Constater que, si $\alpha \in \mathbb{R}^{+*}$, alors $H_T^\alpha(\mathbb{R}, \mathbb{K})$ est un sous-espace vectoriel de $C_T^0(\mathbb{R}, \mathbb{K})$ stable par $\times$.

7/ a) Montrer que, si $(x, y) \in \mathbb{R}^2$, il existe un élément $\tilde{x}$ de $\mathbb{R}$ tel que :

$$\exists k \in \mathbb{Z} \quad \tilde{x} = x + k.T \quad \text{et} \quad |y - \tilde{x}| \leq T \quad \text{et} \quad |y - \tilde{x}| \leq |y - x|.$$

b) En conclure que $H_T^\beta(\mathbb{R}, \mathbb{K})$ contient $H_T^\alpha(\mathbb{R}, \mathbb{K})$ si $(\alpha, \beta) \in \mathbb{R}^{+*2}$ et $\beta \leq \alpha$.

PARTIE III

Dans cette partie, $\alpha \in]0, 1[$ et I_α est l'intégrale impropre :

$$\int_0^{+\infty} (u^{\alpha-2} \cdot (1 - \cos u)) \cdot du$$

8/ Justifier qu'il soit licite de considérer une application de $\mathbb{R}$ dans $\mathbb{R}$, notée f_α , continue, T -périodique, paire et satisfaisant :

$$\forall t \in]0, T/2] \quad f_\alpha(t) = t^\alpha$$

9/ a) Vérifier que : $\forall t \in]1, +\infty[\quad t^\alpha - 1 \leq (t - 1)^\alpha$.

b) Prouver que $H^\alpha([0, T], \mathbb{K})$ contient la restriction de f_α à $[0, T]$.

c) En déduire que $f_\alpha \in H_T^\alpha(\mathbb{R}, \mathbb{R})$. Appartient-elle à $H_T^\beta(\mathbb{R}, \mathbb{R})$ si $\beta \in]\alpha, +\infty[$?

10/ Montrer que I_α converge puis que $I_\alpha > 0$.

11/ a) Déterminer, en fonction de I_α , un équivalent de $c_n(f)$ lorsque n tend vers $+\infty$.

b) $(c_n(f))_{n \in \mathbb{Z}}$ appartient-elle à F ?

PARTIE IV

Dans toute la suite du problème, $(a, b) \in \mathbb{R}^2$ et $a < b$ et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. De plus, $VB([a, b], \mathbb{K})$ désigne l'ensemble des applications à variations bornées de $[a, b]$ dans $\mathbb{K}$.

12/ a) Constater que, si $n \in \mathbb{N}^*$, alors le uplet $(x_k)_{k \in \llbracket 0, n+1 \rrbracket}$ défini par

$$x_0 = 0, x_{n+1} = 1 \text{ et } \forall k \in \llbracket 1, n \rrbracket \quad x_k = \frac{1}{2 \cdot (n+1-k)}$$

est une subdivision de $[0, 1]$.

b) Montrer que l'application φ de $]0, 1[$ dans $\mathbb{R}$ définie par $\varphi(x) = x \cdot \cos\left(\frac{\pi}{2 \cdot x}\right)$ admet un prolongement à $[0, 1]$ continu puis que celui-ci n'est pas à variation bornée.

13/ a) Constater que toute application monotone de $[a, b]$ dans $\mathbb{R}$ est à variation bornée et préciser sa variation.

b) Prouver que $VB([a, b], \mathbb{K})$ contient $H^1([a, b], \mathbb{K})$. Qu'en déduit-on sur toute application de $[a, b]$ dans $\mathbb{K}$ de classe C^1 ?

14/ a) Vérifier que $VB([a, b], \mathbb{K})$ est un $\mathbb{K}$ -espace vectoriel.

b) Soient $f \in VB([a, b], \mathbb{K})$ et $(\lambda, \mu) \in \mathbb{R}^* \times \mathbb{R}$. Démontrer que la fonction $\tilde{f}$ définie par $\tilde{f}(t) = f(\lambda \cdot t + \mu)$ est à variation bornée sur son domaine de définition. Comparer les variations de f et $\tilde{f}$.

15/ Soient $f \in \mathbb{K}^{[a, b]}$ et $c \in]a, b[$. Justifier que f est à variation bornée si, et seulement si, les restrictions de f à $[a, c]$ et $[c, b]$ le sont. Montrer que, sous cette hypothèse :

$$V_{[a, b]}(f) = V_{[a, c]}(f) + V_{[c, b]}(f).$$

16/ Soit $f \in C_T^0(\mathbb{R}, \mathbb{K})$ telle que la restriction de f à $[0, T]$ soit à variation bornée. On considère $K \in \mathbb{N}^*$ et une subdivision $(x_k)_{k \in \llbracket 0, K \rrbracket}$ de $[0, T]$.

a) Vérifier que, si $n \in \mathbb{N}^*$, alors :

$$\left| \sum_{k=1}^K \int_{x_{k-1}}^{x_k} \left[(f(t) - f(x_k)) \cdot e^{-i \cdot n \cdot \omega \cdot t} \right] \cdot dt \right| \leq \sum_{k=1}^K \left[V_{[x_{k-1}, x_k]}(f) \cdot (x_k - x_{k-1}) \right]$$

puis que :

$$\left| \sum_{k=1}^K \int_{x_{k-1}}^{x_k} \left[f(x_k) \cdot e^{-i \cdot n \cdot 2 \cdot \pi \cdot t / T} \right] \cdot dt \right| \leq \frac{V_{[0, T]}(f)}{|n| \cdot \omega}$$

b) En conclure que : $\forall n \in \mathbb{N}^*, |c_n(f)| \leq \frac{V_{[0, T]}(f)}{2 \cdot \pi \cdot |n|}$. Qu'en déduit-on sur $(c_n)_{n \in \mathbb{Z}}$?

PARTIE V

Dans cette dernière partie, $\alpha \in]0, 1[$ et $p \in]1, 1/\alpha[$. On se propose de prouver que $H^\alpha([a, b], \mathbb{K})$ n'est pas inclus dans $VB([a, b], \mathbb{K})$.

17/ a) Exhiber un nombre réel ρ tel que la suite $(\lambda_n)_{n \in \mathbb{N}}$ définie par

$$\forall n \in \mathbb{N} \quad \lambda_n = 1 - \rho \cdot \sum_{k=1}^n k^{-p}$$

converge vers 0.

b) Vérifier qu'alors elle est à valeurs dans $]0, 1]$ et qu'elle décroît.

18/ Si $n \in \mathbb{N}$, on pose $\mu_n = \frac{\lambda_n + \lambda_{n+1}}{2}$ et on note g_n l'application de $[0, 1]$ dans $\mathbb{R}$ définie par

$$\forall t \in [0, 1] \quad g_n(t) = \max \left\{ 0, \frac{\lambda_n - \lambda_{n+1}}{2} - |t - \mu_n| \right\}$$

a) Représenter graphiquement (sans justification) l'application g_n , montrer que $g_n \in H^1([0, 1], \mathbb{R})$ puis évaluer $V_{[0,1]}(g_n)$, lorsque $n \in \mathbb{N}$.

b) Montrer qu'il est légitime de considérer l'application h définie par :

$$\forall t \in [0, 1] \quad h(t) = \sum_{k=1}^{+\infty} g_k(t)$$

puis que $h \in H^1([0, 1], \mathbb{R})$.

c) Montrer que $h^\alpha \in H^\alpha([0, 1], \mathbb{R})$.

19/ a) Evaluer $V_{[\lambda_n, 1]}(h^\alpha)$ lorsque $n \in \mathbb{N}$.

b) En déduire que h^α n'appartient pas à $VB([0, 1], \mathbb{R})$.

20/ a) Conclure.

b) A quelle condition, portant sur l'élément β de $\mathbb{R}^{*+}$, $VB([a, b], \mathbb{K})$ contient-il $H^\beta([a, b], \mathbb{K})$?

Fin

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2014

Épreuve à option (A) : Mathématiques

Durée : 4h

Le problème a pour objet l'extension de la notion de série entière à des séries de matrices, dans le but de définir le logarithme d'une matrice au voisinage de la matrice-identité et d'en examiner certaines propriétés.

Les quatre parties du problème sont largement indépendantes. Quel que soit l'ordre dans lequel le candidat décide de les aborder, on attend de lui une démarche claire, ainsi qu'une rédaction soignée. La présentation, la précision des arguments et la rigueur des raisonnements seront prises en compte dans la notation.

Dans tout le problème, p désigne un nombre entier strictement positif. On note $\mathcal{M}_p(\mathbb{C})$ l'ensemble des matrices carrées d'ordre p à coefficients complexes, et $\mathcal{M}_p(\mathbb{R})$ l'ensemble des matrices de $\mathcal{M}_p(\mathbb{C})$ dont tous les coefficients sont réels.

Pour toute matrice M de $\mathcal{M}_p(\mathbb{C})$ et tout couple (ℓ, k) d'entiers compris entre 1 et p , on note $M[\ell, k]$ le coefficient situé dans la ℓ -ième ligne et la k -ième colonne de la matrice M .

Dans la base des matrices élémentaires $E_{\ell, k}$ (matrice dont tous les coefficients valent 0, sauf celui situé dans la ℓ -ième ligne et la k -ième colonne, qui vaut 1) une matrice M de $\mathcal{M}_p(\mathbb{C})$ s'écrit donc :

$$M = \sum_{\ell=1}^p \sum_{k=1}^p M[\ell, k] E_{\ell, k}.$$

On dit qu'une matrice M de $\mathcal{M}_p(\mathbb{C})$ est triangulaire si tous ses coefficients sous-diagonaux (les $M[\ell, k]$ pour lesquels $\ell > k$) sont nuls.

On appelle polynôme caractéristique de M le polynôme $\chi_M = \det(X.I_p - M)$, où I_p désigne la matrice-identité d'ordre p . Pour tout entier k compris entre 0 et p , on notera $a_k(M)$ le coefficient de son monôme de degré k , de sorte que :

$$\chi_M = \sum_{k=0}^p a_k(M) X^k, \text{ avec } a_p(M) = 1 \text{ et } a_0(M) = (-1)^p \det(M).$$

Le spectre de M , constitué des racines (complexes) de χ_M , sera noté $\text{sp}(M)$:

$$\text{sp}(M) = \{\lambda \in \mathbb{C}; \det(\lambda.I_p - M) = 0\}.$$

On dit qu'une suite $(M_n)_{n \in \mathbb{N}}$ de matrices de $\mathcal{M}_p(\mathbb{C})$ est convergente si chacune des suites $(M_n[\ell, k])_{n \in \mathbb{N}}$ converge dans $\mathbb{C}$.

On dit que la série de terme général M_n est convergente si la suite $(\sum_{m=0}^n M_m)_{n \in \mathbb{N}}$ est convergente.

Partie 1 : convergence d'une série entière de matrices

Soit $\| \cdot \|$ une norme sur l'espace vectoriel $\mathcal{M}_{p,1}(\mathbb{C})$ des matrices-colonnes à p coefficients complexes.

On note $\| \cdot \|_s$ la norme sur $\mathcal{M}_p(\mathbb{C})$ subordonnée à $\| \cdot \|$, définie par :

$$\forall A \in \mathcal{M}_p(\mathbb{C}), \| A \|_s = \text{Sup}\{ \| AY \|; Y \in \mathcal{M}_{p,1}(\mathbb{C}), \| Y \| \leq 1 \}.$$

1. Justifier pour toute matrice Y de $\mathcal{M}_{p,1}(\mathbb{C})$ et tout couple (A, B) de matrices de $\mathcal{M}_p(\mathbb{C})$ les inégalités :

$$\|AY\| \leq \|A\|_s \|Y\| ,$$

$$\|AB\|_s \leq \|A\|_s \|B\|_s .$$

2. Pour toute matrice $A \in \mathcal{M}_p(\mathbb{C})$, on appelle rayon spectral de A , noté $\rho(A)$, le plus grand des modules des valeurs propres (complexes) de A : $\rho(A) = \max\{|\lambda|; \lambda \in \text{sp}(A)\}$.
Démontrer que, pour toute matrice $A \in \mathcal{M}_p(\mathbb{C})$, on a : $\rho(A) \leq \|A\|_s$.

3. Dans cette question, $(c_n)_{n \in \mathbb{N}}$ désigne une suite de nombres complexes, telle que le rayon de convergence de la série entière $\sum_{n \geq 0} c_n z^n$ soit strictement positif.

On note R ce rayon de convergence, éventuellement infini.

On note $\mathcal{A}_c$ l'ensemble des matrices A de $\mathcal{M}_p(\mathbb{C})$ telles que la série $\sum_{n \geq 0} c_n A^n$ soit convergente.

- a) Établir les deux inclusions :

$$\{A \in \mathcal{M}_p(\mathbb{C}); \|A\|_s < R\} \subset \mathcal{A}_c \subset \{A \in \mathcal{M}_p(\mathbb{C}); \rho(A) \leq R\} .$$

- b) Quel résultat sur les séries entières retrouve-t-on lorsque p est égal à 1, si on identifie alors $\mathcal{M}_p(\mathbb{C})$ à $\mathbb{C}$?

- c) Démontrer que $\mathcal{A}_c$ est invariant par similitude, c'est-à-dire que, si A appartient à $\mathcal{A}_c$ et si A' est semblable à A , alors A' est aussi un élément de $\mathcal{A}_c$.

- d) Démontrer que si A est une matrice diagonalisable de $\mathcal{M}_p(\mathbb{C})$ dont le rayon spectral est strictement inférieur à R , alors A appartient à $\mathcal{A}_c$.

Partie 2 : quelques calculs et un exemple

Pour tout couple $(n, z) \in \mathbb{N} \times \mathbb{C}$, on note $u_n(z) = (-1)^n \frac{z^{n+1}}{n+1}$ et $v_n(z) = \frac{z^{2n+1}}{2n+1}$.

1. a) Démontrer que la suite $d = (d_n)_{n \in \mathbb{N}^*}$ de terme général $d_n = \ln(n) - \sum_{k=1}^n \frac{1}{k}$ est convergente.

- b) Exprimer, pour $N \in \mathbb{N}^*$, la somme partielle $\sum_{n=0}^{2N-1} u_n(1)$ en fonction de d_{2N} et d_N ,
et en déduire l'égalité : $\sum_{n=0}^{\infty} u_n(1) = \ln(2)$.

- c) Pour quelles valeurs réelles de x la série de terme général $u_n(x)$ est-elle convergente ? Quelle est alors sa somme ?
2. a) Préciser le rayon de convergence de la série entière de terme général v_n .
- b) Calculer $\sum_{n=0}^{\infty} v_n(x)$, pour tout réel x tel que la série $\sum_{n=0}^{\infty} v_n(x)$ converge.
- c) Établir, pour tout réel $x \in [-1, +1]$, l'égalité : $\sum_{n=0}^{\infty} v_n(ix) = i \arctan(x)$.
- d) Justifier la convergence de la série $\sum_{n=0}^{\infty} u_n(i)$ et calculer sa somme.
3. Dans cette question, on note : $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
- a) Justifier l'existence d'une matrice inversible P de $\mathcal{M}_2(\mathbb{C})$ telle que la matrice $D = P^{-1}AP$ soit diagonale.
- b) Utiliser le résultat de la question 2d pour calculer $\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} D^{n+1}$.
- c) En déduire $\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} A^{n+1}$.

Partie 3 : quelques questions topologiques

1. Soit T une matrice triangulaire de $\mathcal{M}_p(\mathbb{C})$.
- Pour tout entier strictement positif n on pose : $T_n = T + \sum_{\ell=1}^p \frac{\ell}{n} E_{\ell,\ell}$.
- a) Calculer la limite, quand n tend vers l'infini, de la différence $T_n[\ell, \ell] - T_n[k, k]$ de deux coefficients diagonaux de T_n .
- b) Démontrer qu'il existe un entier N tel que, pour tout $n \geq N$, les coefficients diagonaux de T_n soient deux à deux distincts.
2. a) Rappeler pourquoi toute matrice de $\mathcal{M}_p(\mathbb{C})$ est trigonalisable.
- b) Démontrer que l'ensemble des matrices diagonalisables de $\mathcal{M}_p(\mathbb{C})$ est dense dans $\mathcal{M}_p(\mathbb{C})$.
3. Pour tout polynôme $Q = \sum_{k=0}^p b_k X^k$ de $\mathbb{C}_p[X]$, on note : $\|Q\|_1 = \sum_{k=0}^p |b_k|$.
- Justifier, pour toute matrice $M \in \mathcal{M}_p(\mathbb{C})$, l'existence d'un réel $\alpha > 0$ tel que :

$$\forall N \in \mathcal{M}_p(\mathbb{C}), \quad \|N - M\|_s \leq \alpha \Rightarrow \|\chi_N - \chi_M\|_1 \leq 1.$$

4. Soit $M \in \mathcal{M}_p(\mathbb{C})$ et $\mu \in \text{sp}(M)$.

a) Justifier successivement les inégalités :

$$(1) |\mu|^p \leq \text{Max}\{1, |\mu|^{p-1}\} \sum_{k=0}^{p-1} |a_k(M)| \quad (2) \text{Max}\{1, |\mu|^p\} \leq \text{Max}\{1, |\mu|^{p-1}\} \sum_{k=0}^p |a_k(M)|$$

et (3) $|\mu| \leq \|\chi_M\|_1$.

b) En déduire que, pour toute matrice $N \in \mathcal{M}_p(\mathbb{C})$ telle que $\|\chi_N - \chi_M\|_1 \leq 1$, on

a :

$$\forall \lambda \in \text{sp}(N), \quad |\lambda| \leq 2 \|\chi_M\|_1.$$

5. Soit $(M_n)_{n \in \mathbb{N}}$ une suite convergente de matrices de $\mathcal{M}_p(\mathbb{R})$, de limite M .

On suppose que, pour tout $n \in \mathbb{N}$, le polynôme caractéristique de M_n est scindé dans $\mathbb{R}[X]$ c'est-à-dire qu'il admet p racines réelles $\lambda_{1,n}, \dots, \lambda_{p,n}$, distinctes ou non :

$$\chi_{M_n} = \prod_{k=1}^p (X - \lambda_{k,n}).$$

a) Justifier l'existence d'une suite convergente extraite de la suite $((\lambda_{1,n}, \dots, \lambda_{p,n}))_{n \in \mathbb{N}}$.

b) En déduire que le polynôme caractéristique de M est scindé dans $\mathbb{R}[X]$.

6. Soit $\mathcal{D}_p(\mathbb{R})$ l'ensemble des matrices M de $\mathcal{M}_p(\mathbb{R})$ qui sont diagonalisables dans $\mathbb{R}$, c'est-à-dire telles qu'il existe une matrice inversible P de $\mathcal{M}_p(\mathbb{R})$ pour laquelle la matrice $P^{-1}MP$ est diagonale.

a) Démontrer que l'adhérence de $\mathcal{D}_p(\mathbb{R})$ est l'ensemble des matrices de $\mathcal{M}_p(\mathbb{R})$ dont le spectre est inclus dans $\mathbb{R}$.

b) Montrer que si le polynôme caractéristique d'une matrice M de $\mathcal{D}_p(\mathbb{R})$ n'est pas scindé à racines simples, alors il existe une suite de matrices non diagonalisables convergeant vers M .

c) Que peut-on en déduire sur l'intérieur de $\mathcal{D}_p(\mathbb{R})$?

Partie 4 : logarithme de matrice

On note $\mathcal{T}_p(\mathbb{R})$ l'ensemble des matrices M de $\mathcal{M}_p(\mathbb{R})$ qui sont trigonalisables dans $\mathbb{R}$, c'est-à-dire telles qu'il existe une matrice inversible P de $\mathcal{M}_p(\mathbb{R})$ pour laquelle la matrice $P^{-1}MP$ est triangulaire.

Soit $\mathcal{B} = \{M \in \mathcal{T}_p(\mathbb{R}); \|M - I_p\|_s < 1\}$.

1. Soit $M \in \mathcal{B}$.

- a) Justifier la convergence de la série $\sum_{n \geq 0} \frac{(-1)^n}{n+1} (M - I_p)^{n+1}$.
- b) Démontrer que $\text{sp}(M)$ est inclus dans l'intervalle ouvert $]0, 2[$.

Pour toute matrice $M \in \mathcal{B}$, on note désormais : $\text{Log}(M) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} (M - I_p)^{n+1}$.

2. Démontrer que, pour toute matrice $M \in \mathcal{B}$, la trace de $\text{Log}(M)$ est égale à $\ln(\det M)$ (on pourra commencer par le cas où M est diagonalisable).

3. Soit M et N deux matrices de $\mathcal{T}_p(\mathbb{R})$ telles que : $MN = NM$.

a) Démontrer qu'il existe une matrice inversible P de $\mathcal{M}_p(\mathbb{R})$ pour laquelle les deux matrices $P^{-1}MP$ et $P^{-1}NP$ sont triangulaires.

b) Démontrer qu'il existe deux suites $(M_n)_{n \in \mathbb{N}}$ et $(N_n)_{n \in \mathbb{N}}$ de matrices de $\mathcal{D}_p(\mathbb{R})$ qui convergent respectivement vers M et N , telles que, pour tout $n \in \mathbb{N}$, on ait :

$$M_n N_n = N_n M_n.$$

4. Utiliser le résultat précédent pour prouver que, si M et N sont deux matrices de $\mathcal{B}$ telles que $MN = NM$ et $MN \in \mathcal{B}$, alors :

$$\text{Log}(MN) = \text{Log}(M) + \text{Log}(N).$$

**Banque d'Épreuves des Concours des Écoles
d'Actuariat et Statistique**

Session 2014

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet est constitué d'un problème en 4 parties et d'un exercice indépendant du problème. On pourra choisir de commencer par traiter l'exercice.

Rappel de résultats

On rappelle les résultats suivants, qui pourront être utilisés dans la suite sans qu'il soit nécessaire de les redémontrer.

1. **Inégalité de Markov** : pour toute variable aléatoire positive X possédant une espérance, et pour tout nombre réel $x > 0$, on a l'inégalité :

$$\mathbb{P}(X \geq x) \leq \frac{\mathbb{E}(X)}{x}.$$

2. **Inégalité de Bienaymé-Tchebychev** : pour toute variable aléatoire réelle X possédant une espérance et une variance, et pour tout nombre réel $x > 0$, on a l'inégalité :

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq x) \leq \frac{\mathbb{V}(X)}{x^2}.$$

3. **Identité pour l'espérance** : pour toute variable aléatoire positive X , on a

$$\mathbb{E}(X) = \int_0^{+\infty} \mathbb{P}(X > x) dx,$$

l'espérance de la variable aléatoire X étant définie si et seulement si l'intégrale ci-dessus est convergente.

4. **Espérance d'un produit** : si X et Y sont des variables aléatoires réelles indépendantes possédant chacune une espérance, alors la variable aléatoire produit XY possède également une espérance, et

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

5. **Loi faible des grands nombres** : on considère une variable aléatoire réelle X possédant une espérance, et, pour tout $N \geq 1$, et tout $\varepsilon > 0$, on pose

$$p(N, \varepsilon) = \mathbb{P}\left(\left|\frac{X_1 + \dots + X_N}{N} - \mathbb{E}(X)\right| \geq \varepsilon\right),$$

où $X_1, \dots, X_N$ désigne une famille de variables aléatoires indépendantes et identiquement distribuées de même loi que X . La loi faible des grands nombres affirme que l'on a alors :

$$\forall \varepsilon > 0, \lim_{N \rightarrow +\infty} p(N, \varepsilon) = 0.$$

Problème

Cadre mathématique

Dans tout ce problème, X désigne une variable aléatoire positive, dont on note F_X la fonction de répartition, N désigne un entier vérifiant la condition $N \geq 1$, et $X_1, \dots, X_N$ désigne une famille de N variables aléatoires indépendantes et identiquement distribuées de même loi que X . On définit alors la variable aléatoire S_N de la manière suivante :

$$S_N = X_1 + \dots + X_N.$$

L'objectif du problème est l'étude de diverses bornes pour des probabilités de la forme

$$\mathbb{P}\left(\frac{S_N}{N} \geq x\right).$$

Contexte actuariel

Aucune connaissance en assurance n'est nécessaire pour traiter ce problème. Cependant, la question étudiée se rattache à une problématique actuarielle concrète, brièvement esquissée ci-après.

Sur une période d'un an, un assureur prend à sa charge les montants des sinistres affectant une population de N assurés. En modélisant les montants individuels des sinistres associés à chacun des assurés par une famille de variables aléatoires $X_1, \dots, X_N$, la somme S_N représente donc le montant total à la charge de l'assureur sur une période d'un an. En supposant que chacun des N assurés verse une prime d'assurance annuelle égale à x , le montant total des primes perçues pour une année est Nx , et l'événement $\left\{\frac{S_N}{N} \geq x\right\}$ correspond à la situation défavorable où le montant des primes perçues est inférieur au montant des sinistres à prendre en charge. La valeur de la prime x est généralement fixée de manière à rendre suffisamment faible la probabilité d'un tel événement, et l'estimation de probabilités du type $\mathbb{P}\left(\frac{S_N}{N} \geq x\right)$ est donc une question pertinente dans le contexte de l'assurance.

Organisation

Les parties II et III peuvent être traitées indépendamment de la partie I, en admettant simplement le résultat de la question 8) et les définitions données dans la partie I. La partie IV peut être traitée indépendamment des parties I, II, III.

Notation exponentielle

Dans tout le problème, on utilise indifféremment les notations e^x et $\exp(x)$ pour désigner l'exponentielle d'un nombre réel x .

Partie préliminaire

1) On suppose que la variable aléatoire X possède une espérance. Montrer, en faisant appel à la loi des grands nombres, que pour tout $x > \mathbb{E}(X)$, on a

$$\lim_{N \rightarrow +\infty} \mathbb{P}\left(\frac{S_N}{N} \geq x\right) = 0.$$

Partie I – Bornes exponentielles : cadre général

Dans toute cette partie, on suppose que la variable aléatoire X vérifie la propriété suivante : il existe deux constantes $a > 0$ et $b > 0$ telles que

$$\forall x \geq 0, \mathbb{P}(X > x) \leq a \exp(-bx). \quad (1)$$

2) Montrer, grâce à l'inégalité (1), que la variable aléatoire X possède une espérance. (On pourra utiliser l'identité pour l'espérance figurant dans la liste de résultats rappelés en page 2/10.)

3) Pour $t \in [0, +\infty[$ et $z \geq 1$, déduire de l'inégalité (1) une majoration de la probabilité $\mathbb{P}(\exp(tX) > z)$. Donner ensuite une majoration très simple de cette probabilité lorsque $z \in [0, 1[$.

4) En déduire que, pour tout $t \in [0, b[$, la variable aléatoire $\exp(tX)$ possède une espérance, et justifier que $\mathbb{E}(\exp(tX)) > 0$.

Définitions : pour tout $t \in [0, b[$, on définit $\Lambda(t)$ par $\Lambda(t) = \ln(\mathbb{E}(\exp(tX)))$.

5) Montrer que, pour tout $t \in [0, b[$, la variable $\exp(tS_N)$ possède une espérance donnée par

$$\mathbb{E}(\exp(tS_N)) = \exp(N\Lambda(t)).$$

Définitions : pour tout $x \geq 0$ et $t \geq 0$, on définit les événements $A(N, x)$ et $B(N, x, t)$ par

$$A(N, x) = \left\{ \frac{S_N}{N} \geq x \right\}, \quad B(N, x, t) = \{ \exp(-Ntx) \exp(tS_N) \geq 1 \}.$$

6) Montrer que, pour tout $x \geq 0$ et $t \geq 0$, on a l'inclusion entre événements

$$A(N, x) \subset B(N, x, t),$$

et en déduire que

$$\mathbb{P}(A(N, x)) \leq \mathbb{P}(B(N, x, t)).$$

7) Montrer que, pour tout $x \geq 0$, et tout $t \in [0, b[$,

$$\mathbb{P}(B(N, x, t)) \leq \exp(-N(tx - \Lambda(t))).$$

8) Montrer que, pour tout $x \geq 0$, on a l'inégalité :

$$\mathbb{P}\left(\frac{S_N}{N} \geq x\right) \leq \exp(-NI(x)), \quad (2)$$

où $I(x)$ désigne la borne supérieure de l'ensemble des valeurs prises par $tx - \Lambda(t)$ lorsque t décrit l'intervalle $[0, b[$, autrement dit :

$$I(x) = \sup \{ tx - \Lambda(t) \mid t \in [0, b[\}.$$

9) Au vu du résultat de la question 8), est-il possible d'avoir $I(x) > 0$ pour un nombre $x \in [0, \mathbb{E}(X)[$? (Justifier la réponse.)

Partie II – Bornes exponentielles : le cas de la loi exponentielle

Dans cette partie, on suppose que la variable aléatoire X suit la loi exponentielle $\mathcal{E}(\lambda)$, où λ est un nombre réel strictement positif. On rappelle que cette loi est caractérisée par la densité f_X définie sur $\mathbb{R}$ par :

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{si } x \geq 0, \\ 0 & \text{si } x < 0. \end{cases}$$

10) Calculer la valeur de $\mathbb{P}(X > x)$ pour tout $x \geq 0$.

11) En déduire que l'inégalité (1) est satisfaite avec $a = 1$ et $b = \lambda$.

12) Pour $z \geq 0$ et $t \in [0, +\infty[$, calculer la valeur de $\mathbb{P}(\exp(tX) > z)$ en distinguant les cas $z \in [0, 1[$ et $z \in [1, +\infty[$.

13) En déduire que, pour tout $t \in [0, \lambda[$, la variable aléatoire $\exp(tX)$ possède une espérance, et calculer la valeur de celle-ci.

14) En déduire, à l'aide de la question 8), que, pour tout $x \geq \frac{1}{\lambda}$, on a l'inégalité :

$$\mathbb{P}\left(\frac{S_N}{N} \geq x\right) \leq \exp(-NI_{\mathcal{E}(\lambda)}(x)),$$

où l'on a posé :

$$I_{\mathcal{E}(\lambda)}(x) = \lambda x - 1 - \ln(\lambda x).$$

15) Montrer que $I_{\mathcal{E}(\lambda)}(x) > 0$ pour tout $x > \frac{1}{\lambda}$. Que vaut $I_{\mathcal{E}(\lambda)}(x)$ lorsque $x = \frac{1}{\lambda}$?

Partie III – Bornes exponentielles : le cas de la loi de Bernoulli

Dans cette partie, on suppose que X suit une loi de Bernoulli $\mathcal{B}(p)$ de paramètre $p \in]0, 1[$. Autrement dit, la loi de X est caractérisée par :

$$\mathbb{P}(X = x) = \begin{cases} p & \text{si } x = 1, \\ 1 - p & \text{si } x = 0. \end{cases}$$

16) Calculer la valeur de $\mathbb{P}(X > x)$ pour tout $x \geq 0$.

17) Montrer que, pour tout $b > 0$, on peut trouver un $a > 0$ tel que l'inégalité (1) soit satisfaite.

18) Calculer $\mathbb{E}(X)$.

19) Calculer, pour tout $t \geq 0$, la valeur de $\Lambda(t)$.

20) En déduire à l'aide de la question 8) l'inégalité, valable pour tout $x \in]p, 1[$,

$$\mathbb{P}\left(\frac{S_N}{N} \geq x\right) \leq I_{\mathcal{B}(p)}(x),$$

où

$$I_{\mathcal{B}(p)}(x) = x \ln(x/p) + (1-x) \ln\left(\frac{1-x}{1-p}\right).$$

21) Montrer que $I_{\mathcal{B}(p)}(x) > 0$ pour tout $x \in]p, 1[$. Comment se comporte $I_{\mathcal{B}(p)}(x)$ lorsque $x \searrow p$ et lorsque $x \nearrow 1$?

Partie IV - Un exemple de comportement polynomial

Dans cette partie, on suppose que la loi de X est caractérisée par une densité f_X dépendant d'un paramètre $\alpha \in]1, 2[$, et possédant l'expression suivante :

$$f_X(x) = \begin{cases} \alpha(x+1)^{-(\alpha+1)} & \text{si } x \geq 0, \\ 0 & \text{si } x < 0. \end{cases}$$

22) Montrer que la fonction de répartition F_X de la variable aléatoire X est donnée, pour tout $x \geq 0$, par la formule :

$$F_X(x) = 1 - \frac{1}{(x+1)^\alpha}.$$

23) Montrer que la variable aléatoire X possède une espérance, et calculer la valeur de celle-ci.

Définitions : pour $u \geq 0$, on définit la variable aléatoire $W_N(u)$ comme étant le nombre d'indices i entre 1 et N pour lesquels $X_i \geq u$. Autrement dit :

$$W_N(u) = \text{nombre d'éléments de l'ensemble } \{i \in \{1, \dots, N\} \mid X_i \geq u\}.$$

24) Quelle est la loi de $W_N(u)$? En déduire une expression simple de

$$\mathbb{P}(W_N(u) \geq 1).$$

25) Montrer que, pour tout $x > 0$ fixé, on a l'équivalent suivant lorsque $N \rightarrow +\infty$:

$$\mathbb{P}(W_N(xN) \geq 1) \sim x^{-\alpha} N^{1-\alpha}.$$

26) En déduire que, pour tout $x > 0$, et tout $\epsilon \in]0, 1[$, il existe un entier $N_0 \geq 1$ tel que, pour tout $N \geq N_0$,

$$\mathbb{P}\left(\frac{S_N}{N} \geq x\right) \geq (1-\epsilon)x^{-\alpha} N^{1-\alpha}.$$

27) Une inégalité semblable à l'inégalité (2) obtenue à la question 8), dans laquelle on aurait $I(x) > 0$, est-elle possible dans le cas étudié ici ? (Justifier la réponse.)

Définitions : pour tout $u \geq 0$, on introduit la variable aléatoire $Y^{(u)}$ définie par

$$Y^{(u)} = \min(X, u).$$

Autrement dit,

$$Y^{(u)} = \begin{cases} X & \text{lorsque } X \leq u, \\ u & \text{lorsque } X > u. \end{cases}$$

28) Pour $u \geq 0$, calculer la fonction de répartition de la variable aléatoire $Y^{(u)}$.

29) Montrer que, pour tout $u \geq 0$, la variable aléatoire $Y^{(u)}$ possède une espérance donnée par :

$$\mathbb{E}(Y^{(u)}) = \frac{1}{\alpha - 1} \left(1 - \frac{1}{(1 + u)^{\alpha - 1}} \right).$$

30) Montrer que, pour tout $u \geq 0$, la variable aléatoire $[Y^{(u)}]^2 = Y^{(u)} \times Y^{(u)}$, possède une espérance donnée par :

$$\mathbb{E}([Y^{(u)}]^2) = \frac{2}{2 - \alpha} (1 + u)^{2 - \alpha} - \frac{2}{2 - \alpha} - \frac{2}{\alpha - 1} \left(1 - \frac{1}{(1 + u)^{\alpha - 1}} \right).$$

31) En déduire que, pour tout $u \geq 0$, la variable aléatoire $Y^{(u)}$ possède une variance, et donner une expression de celle-ci.

Définitions : pour tout $u \geq 0$, et pour tout $i \in \{1, \dots, N\}$, on définit la variable aléatoire $Y_i^{(u)}$ par $Y_i^{(u)} = \min(X_i, u)$, puis l'on définit la variable aléatoire $R_N^{(u)}$ en posant

$$R_N^{(u)} = Y_1^{(u)} + \dots + Y_N^{(u)}.$$

32) Montrer que la variable aléatoire $\frac{R_N^{(u)}}{N}$ possède une espérance et une variance, respectivement données par

$$\mathbb{E}\left(\frac{R_N^{(u)}}{N}\right) = \mathbb{E}(Y^{(u)}), \quad \mathbb{V}\left(\frac{R_N^{(u)}}{N}\right) = \frac{\mathbb{V}(Y^{(u)})}{N}.$$

33) Montrer que, pour tout $u \geq 0$, et pour tout $x > \mathbb{E}(Y^{(u)})$, on a l'inégalité :

$$\mathbb{P}\left(\frac{R_N^{(u)}}{N} \geq x\right) \leq \frac{\mathbb{V}(Y^{(u)})}{N(x - \mathbb{E}(Y^{(u)}))^2}.$$

34) Montrer que, pour tout $u \geq 0$, on a l'inclusion entre événements :

$$\{W_N(u) = 0\} \subset \{S_N = R_N^{(u)}\}.$$

35) Montrer que, pour tout $u \geq 0$, et pour tout $x > \mathbb{E}(Y^{(u)})$, on a l'inclusion entre événements :

$$\left\{ \frac{S_N}{N} \geq x \right\} \subset \{W_N(u) \geq 1\} \cup \left\{ \frac{R_N^{(u)}}{N} \geq x \right\}.$$

En déduire que

$$\mathbb{P}\left(\frac{S_N}{N} \geq x\right) \leq \mathbb{P}(W_N(u) \geq 1) + \mathbb{P}\left(\frac{R_N^{(u)}}{N} \geq x\right).$$

36) On considère à présent un nombre $x > \frac{1}{\alpha-1}$ fixé. En utilisant les résultats précédents avec $u = xN$, montrer que, pour tout $\epsilon \in]0, 1[$, il existe un entier $N_1 \geq 1$ tel que, pour tout $N \geq N_1$, on a :

$$\mathbb{P}\left(\frac{S_N}{N} \geq x\right) \leq (1 + \epsilon) x^{-\alpha} N^{1-\alpha} g(x),$$

où

$$g(x) := 1 + \frac{2}{(2 - \alpha) \left(1 - \frac{1}{(\alpha-1)x}\right)^2}$$

37) Comparer les inégalités obtenues aux questions 26) et 36).

Exercice

Un assureur proposant des contrats d'assurance automobile utilise le modèle suivant. Deux profils de conducteurs sont pris en compte : les "bons conducteurs" (profil 1), et les "mauvais conducteurs" (profil 2). Pour un conducteur de profil donné $i \in \{1, 2\}$, le nombre de sinistres annuel que l'assureur est amené à indemniser est modélisé par une variable aléatoire de loi de Poisson $\mathcal{P}(\lambda_i)$ de paramètre λ_i . La différence entre les bons et les mauvais conducteurs se traduit par l'inégalité :

$$\lambda_1 < \lambda_2. \quad (3)$$

Par ailleurs, pour $i \in \{1, 2\}$, on note ρ_i la probabilité pour qu'un assuré pris au hasard corresponde au profil i , et l'on suppose que $\rho_1 > 0$ et $\rho_2 > 0$.

Dans la suite, on note D la variable aléatoire correspondant au nombre de sinistres annuel pour un assuré pris au hasard, et l'on note Π la variable aléatoire à valeurs dans $\{1, 2\}$ indiquant le profil de ce même assuré.

On rappelle la définition de la loi de Poisson : si X suit la loi de Poisson $\mathcal{P}(\lambda)$ de paramètre $\lambda > 0$, on a, pour tout entier $k \in \mathbb{N}$,

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

On rappelle en outre les formules valables dans ce contexte :

$$\mathbb{E}(X) = \lambda$$

et

$$\mathbb{E}(X^2) = \lambda^2 + \lambda.$$

- A) Pour $k \in \mathbb{N}$, expliciter la valeur de $\mathbb{P}(D = k | \Pi = 1)$ et de $\mathbb{P}(D = k | \Pi = 2)$.
- B) Pour $k \in \mathbb{N}$, exprimer la valeur de $\mathbb{P}(D = k)$ en fonction de $k, \lambda_1, \lambda_2, \rho_1, \rho_2$.
- C) Calculer l'espérance $\mathbb{E}(D)$ en fonction de $\lambda_1, \lambda_2, \rho_1, \rho_2$.
- D) Calculer l'espérance $\mathbb{E}(D^2)$ en fonction de $\lambda_1, \lambda_2, \rho_1, \rho_2$.
- E) En déduire une expression de la variance $\mathbb{V}(D)$ en fonction de $\lambda_1, \lambda_2, \rho_1, \rho_2$.
- F) Comparer $\mathbb{V}(D)$ à la variance d'une variable aléatoire de loi de Poisson de paramètre $\lambda = \rho_1 \lambda_1 + \rho_2 \lambda_2$.
- G) Calculer $\mathbb{P}(\Pi = 1 | D = k)$ et $\mathbb{P}(\Pi = 2 | D = k)$.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2015

Épreuve de mathématiques

Durée : 4h

L'objet du problème est l'étude du nombre de records d'une permutation via la suite de variables aléatoires $(R_n)_{n \in \mathbb{N}^*}$ introduite dans la partie II. La partie III utilise les notations et des résultats de la partie II.

Partie I Une distance entre lois de variables aléatoires

Toutes les variables aléatoires de cette partie sont définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$.

1. Soit X, Y et Z trois variables aléatoires à valeurs dans $\mathbb{N}$.

a) Justifier la convergence de la série de terme général $\left| \mathbf{P}([X = n]) - \mathbf{P}([Y = n]) \right|$.

On note $d(X, Y) = \frac{1}{2} \sum_{n=0}^{+\infty} \left(\left| \mathbf{P}([X = n]) - \mathbf{P}([Y = n]) \right| \right)$.

b) Que dire des variables X et Y quand $d(X, Y) = 0$?

Donner un exemple simple de deux variables X et Y distinctes et telles que $d(X, Y) = 0$.

c) Prouver l'inégalité : $d(X, Z) \leq d(X, Y) + d(Y, Z)$.

2. a) On pose $A = \{k \in \mathbb{N}; \mathbf{P}([X = k]) \geq \mathbf{P}([Y = k])\}$. Vérifier l'égalité :

$$d(X, Y) = \left| \mathbf{P}([X \in A]) - \mathbf{P}([Y \in A]) \right|.$$

b) Soit B une partie de $\mathbb{N}$.

Prouver l'inégalité : $\left| \mathbf{P}([X \in B]) - \mathbf{P}([Y \in B]) \right| \leq d(X, Y)$.

On pourra faire intervenir les parties $B \cap A$ et $B \cap A^c$.

3. Soit $p \in [0, 1]$, X une variable suivant la loi de Bernoulli de paramètre p et Y une variable suivant la loi de Poisson de paramètre p .

a) Justifier, pour tout réel x , l'inégalité : $e^x \geq 1 + x$.

b) Établir l'égalité $d(X, Y) = p(1 - e^{-p})$ et en déduire la majoration $d(X, Y) \leq p^2$.

Dans la fin de cette partie, on considère un entier n au moins égal à 3, un entier $N > n$, et on note I la matrice identité de $M_N(\mathbb{R})$ et R la matrice de $M_N(\mathbb{R})$ dont tous les coefficients

sont nuls sauf ceux de la sur-diagonale qui valent 1, c'est-à-dire $R = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & (0) & \ddots & 1 \\ & & & 0 \end{pmatrix}$.

On considère une n -liste $(p_1, p_2, \dots, p_n)$ de réels de $[0, 1]$, n variables aléatoires $X_1, X_2, \dots, X_n$ (resp. $Y_1, Y_2, \dots, Y_n$), indépendantes suivant des lois de Bernoulli (resp. Poisson) de paramètres respectifs $p_1, p_2, \dots, p_n$.

On note, pour tout entier $i \in \llbracket 1, n \rrbracket$, $P_i = (1 - p_i)I + p_i R$ et $Q_i = p_i(R - I)$.

4. a) Montrer que la première ligne du produit $P_1 P_2$ est constituée des $n+1$ réels $\mathbf{P}([X_1 + X_2 = 0])$, $\mathbf{P}([X_1 + X_2 = 1])$, ..., $\mathbf{P}([X_1 + X_2 = n])$ suivis de termes nuls.

b) On note $U_n = \sum_{k=1}^n X_k$. Montrer que la première ligne du produit $P_1 P_2 \dots P_n$ est constituée des $n+1$ réels $\mathbf{P}([U_n = 0])$, $\mathbf{P}([U_n = 1])$, ..., $\mathbf{P}([U_n = n])$ suivis de termes nuls.

5. On rappelle (ou on admet) que si $A \in M_N(\mathbb{R})$, la suite de terme général $\sum_{k=0}^r \frac{A^k}{k!}$ converge, quand r tend vers $+\infty$, vers une matrice notée $\exp A$. La convergence d'une suite de matrices s'entend comme la convergence coefficient par coefficient. On **admet** que si deux matrices A et B de $M_N(\mathbb{R})$ commutent alors on a $\exp(A + B) = \exp A \times \exp B$.

a) Soit i un entier de $\llbracket 1, n \rrbracket$ et r un entier naturel. Prouver l'égalité :

$$\sum_{k=0}^r \frac{Q_i^k}{k!} = \sum_{j=0}^r \frac{p_i^j}{j!} \left(\sum_{k=0}^{r-j} \frac{(-1)^k p_i^k}{k!} \right) R^j.$$

b) En déduire que $\exp Q_i = \sum_{j=0}^{N-1} \frac{p_i^j e^{-p_i}}{j!} R^j$.

c) On note $V_n = \sum_{k=1}^n Y_k$. Montrer que la première ligne du produit $\exp Q_1 \times \exp Q_2 \times \dots \times \exp Q_n$ est constituée par les N réels $\mathbf{P}([V_n = 0])$, $\mathbf{P}([V_n = 1])$, $\dots$, $\mathbf{P}([V_n = N-1])$.

6. On note, pour toute matrice $A = (a_{i,j})_{1 \leq i,j \leq N}$ de $M_N(\mathbb{R})$, $\|A\| = \max_{1 \leq i \leq N} \sum_{j=1}^N |a_{i,j}|$.

a) Prouver, pour tout couple (A, B) d'éléments de $M_N(\mathbb{R})$ les inégalités

$$\|A + B\| \leq \|A\| + \|B\| \quad \text{et} \quad \|AB\| \leq \|A\| \|B\|.$$

b) Prouver, pour tout entier $i \in \llbracket 1, n \rrbracket$, l'inégalité $\|\exp Q_i\| \leq 1$.

c) En remarquant que

$$\prod_{i=1}^n P_i - \prod_{i=1}^n \exp Q_i = (P_1 - \exp Q_1) \left(\prod_{i=2}^n P_i \right) + (\exp Q_1) \left(\prod_{i=2}^n P_i - \prod_{i=2}^n \exp Q_i \right),$$

prouver l'inégalité

$$\left\| \prod_{i=1}^n P_i - \prod_{i=1}^n \exp Q_i \right\| \leq \sum_{i=1}^n \|P_i - \exp Q_i\|.$$

d) En exprimant, pour tout entier i de $\llbracket 1, n \rrbracket$, la matrice $P_i - \exp Q_i$ à l'aide des puissances de la matrice R , prouver l'inégalité $\|P_i - \exp Q_i\| \leq 2p_i^2$.

7. a) Déduire des questions précédentes l'inégalité $d(U_n, V_n) \leq \sum_{i=1}^n p_i^2$.

b) Soit $\lambda > 0$. En appliquant le résultat précédent, retrouver la propriété d'approximation usuelle de la loi binomiale de paramètres n et $\frac{\lambda}{n}$ par une loi de Poisson.

Partie II Records d'une permutation

On rappelle qu'une permutation d'un ensemble non vide E est une bijection de E sur E . On note, pour tout entier naturel n non nul, $\mathcal{S}_n$ l'ensemble des permutations de l'ensemble $\llbracket 1, n \rrbracket$ (des entiers k tels que $1 \leq k \leq n$) et $\sigma := (\sigma_1, \sigma_2, \dots, \sigma_n)$ la permutation σ qui envoie l'entier 1 sur σ_1 , 2 sur σ_2 , ..., n sur σ_n .

On note $\mathcal{P}(\mathcal{S}_n)$ l'ensemble des parties de $\mathcal{S}_n$ et on munit l'espace probabilisable $(\mathcal{S}_n, \mathcal{P}(\mathcal{S}_n))$ de la probabilité uniforme notée $\mathbf{P}$. Ainsi, pour tout élément σ de $\mathcal{S}_n$, on a $\mathbf{P}(\{\sigma\}) = \frac{1}{n!}$.

Soit $n \in \mathbb{N}^*$ et $k \in \llbracket 1, n \rrbracket$. On dit qu'une permutation $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ présente un record au rang k si, pour tout entier i vérifiant $1 \leq i \leq k$, on a $\sigma_i \leq \sigma_k$. Ainsi, en particulier, toute permutation présente un record au rang 1. On note $R_n(\sigma)$ le nombre de records que compte la permutation σ . On définit ainsi une variable aléatoire R_n sur $\mathcal{S}_n$. Bien sûr, la variable R_1 est certaine égale à 1. On note, pour tout entier naturel n non nul, $H_n = \sum_{k=1}^n \frac{1}{k}$.

1. Montrer que la suite $(H_n - \ln n)_{n \in \mathbb{N}^*}$ est monotone convergente. On note γ la limite de cette suite.

Dans toute la suite de cette partie II on considère un entier n au moins égal à 3.

2. Déterminer la loi de R_3 , son espérance et sa variance.

3. Déterminer les probabilités $\mathbf{P}([R_n = 1])$ et $\mathbf{P}([R_n = n])$.

4. a) Soit p un entier de $\llbracket 2, n \rrbracket$. Déterminer le nombre de permutations de $\llbracket 1, n \rrbracket$ ayant exactement 2 records lesquels sont atteints aux rangs 1 et p .

b) Prouver l'égalité : $\mathbf{P}([R_n = 2]) = \frac{1}{n} \sum_{k=1}^{n-1} \frac{1}{k}$.

c) Donner un équivalent de $\mathbf{P}([R_n = 2])$ quand n tend vers l'infini.

5. On introduit, pour tout entier i de $\llbracket 1, n \rrbracket$, la variable aléatoire T_i , définie sur $(\mathcal{S}_n, \mathcal{P}(\mathcal{S}_n), \mathbf{P})$, qui, à chaque élément σ de $\mathcal{S}_n$, associe la valeur 1 si σ présente un record au rang i et égale à 0 sinon.

a) Montrer que, pour tout entier i de $\llbracket 1, n \rrbracket$, T_i suit la loi de Bernoulli de paramètre $\frac{1}{i}$.

b) Calculer l'espérance de la variable R_n et en donner un équivalent simple quand n tend vers l'infini.

6. a) Soit (i, j) un couple d'entiers vérifiant $2 \leq i < j \leq n$.

En calculant la probabilité $\mathbf{P}([T_i = 1] \cap [T_j = 1])$, justifier l'indépendance des variables T_i et T_j .

b) Calculer la variance de la variable R_n et en donner un équivalent simple quand n tend vers l'infini.

7. On **admet** l'indépendance mutuelle des variables $T_1, T_2, \dots, T_n$.

Établir, pour tout entier k de $\llbracket 2, n \rrbracket$, l'égalité :

$$\mathbf{P}([R_n = k]) = \sum_{2 \leq i_2 < i_3 < \dots < i_k \leq n} \frac{1}{i_2} \frac{1}{i_3} \dots \frac{1}{i_k} \prod_{\substack{2 \leq j \leq n \\ j \notin \{i_2, \dots, i_k\}}} (1 - \frac{1}{j}).$$

a) En déduire l'égalité : $\mathbf{P}([R_n = 3]) = \frac{1}{n} \sum_{1 \leq i < j \leq n-1} \frac{1}{ij}$.

b) Prouver l'équivalence : $\mathbf{P}([R_n = 3]) \sim \frac{1}{2} \frac{\ln^2 n}{n}$.

Partie III Deux résultats asymptotiques

1. Un premier résultat

a) Soit $\varepsilon > 0$. Prouver, pour tout entier n assez grand, l'inclusion entre événements :

$$\left[\left| \frac{R_n}{\ln n} - 1 \right| \geq \varepsilon \right] \subset \left[\left| \frac{R_n}{\ln n} - \frac{H_n}{\ln n} \right| \geq \frac{\varepsilon}{2} \right].$$

b) Soit $\varepsilon > 0$.

(i) Établir l'égalité : $\lim_{n \rightarrow +\infty} \mathbf{P} \left(\left| \frac{R_n}{\ln n} - \frac{H_n}{\ln n} \right| \geq \varepsilon \right) = 0$.

(ii) En déduire l'égalité : $\lim_{n \rightarrow +\infty} \mathbf{P} \left(\left| \frac{R_n}{\ln n} - 1 \right| \geq \varepsilon \right) = 0$.

2. Un second résultat

On rappelle que si X est une variable aléatoire à valeurs dans $\mathbb{N}$ on note G_X sa fonction génératrice c'est-à-dire la fonction qui, à chaque réel t de $[0, 1]$, associe

$$G_X(t) = \mathbf{E}(t^X) = \sum_{k=0}^{+\infty} \mathbf{P}([X = k]) t^k.$$

a) Déterminer la fonction génératrice d'une variable de Bernoulli de paramètre $p \in [0, 1]$.

b) Déterminer la fonction génératrice d'une variable de Poisson de paramètre $\lambda > 0$.

c) Soit $n \in \mathbb{N}^*$, et $X_1, X_2, \dots, X_n$ des variables aléatoires, à valeurs dans $\mathbb{N}$, mutuellement indépendantes et $S_n = \sum_{k=1}^n X_k$. Justifier l'égalité : $G_{S_n} = \prod_{k=1}^n G_{X_k}$.

d) Soit $Y, X_1, X_2, \dots$ une suite de variables aléatoires à valeurs dans $\mathbb{N}$. On **admet** le résultat suivant : si, pour tout réel t de $[0, 1]$, la suite $(G_{X_n}(t))_{n \in \mathbb{N}^*}$ converge vers $G_Y(t)$, alors la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers Y , c'est-à-dire que, pour tout entier naturel k ,

$$\lim_{n \rightarrow +\infty} \mathbf{P}([X_n = k]) = \mathbf{P}([Y = k]).$$

Soit m un entier naturel au moins égal à 2 et $n = 2m$. On conserve les notations de la partie II et on pose $W_n = \sum_{k=m+1}^{2m} T_k$ (qui compte le nombre aléatoire de records que présente une permutation de $\llbracket 1, 2m \rrbracket$ entre les rangs $m+1$ et $2m$).

Prouver, pour tout réel t de $[0, 1]$, l'égalité : $G_{W_n}(t) = \prod_{i=m+1}^{2m} \left(1 + \frac{t-1}{i}\right)$.

e) En déduire que, lorsque l'entier m tend vers $+\infty$, la suite de terme général W_n converge en loi vers une variable aléatoire qu'on identifiera.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2015

Épreuve à option (A) : Mathématiques

Durée : 4h

Toutes les matrices de l'énoncé sont à coefficients réels et, pour tout entier n strictement positif, on note $\mathcal{M}_n$ le $\mathbb{R}$ -espace vectoriel des matrices carrées d'ordre n et $\mathcal{C}_n$ le $\mathbb{R}$ -espace vectoriel des matrices-colonnes à n coefficients ($\mathcal{M}_n$ et $\mathcal{C}_n$ sont donc des notations simplifiées pour $\mathcal{M}_n(\mathbb{R})$ et $\mathcal{M}_{n,1}(\mathbb{R})$).

L'espace $\mathcal{M}_1$ est identifié à $\mathbb{R}$ et la transposée d'une matrice M , quel que soit son format, est notée tM .

On s'intéresse dans ce problème aux matrices carrées H dont les coefficients $h_{i,j}$ ne dépendent que de la somme $i + j$ de leur indice de ligne i et de leur indice de colonne j .

Pour toute application $\varphi : \mathbb{N} \rightarrow \mathbb{R}$ et tout entier strictement positif n , on note $H_\varphi^{(n)}$ la matrice carrée de $\mathcal{M}_n$ donnée par :

$$H_\varphi^{(n)} = \left(\varphi(i + j - 2) \right)_{1 \leq i, j \leq n} = \begin{pmatrix} \varphi(0) & \varphi(1) & \dots & \varphi(n-1) \\ \varphi(1) & \varphi(2) & \dots & \varphi(n) \\ \vdots & \vdots & & \vdots \\ \varphi(n-1) & \varphi(n) & \dots & \varphi(2n-2) \end{pmatrix}.$$

L'objet du problème est de préciser dans quels cas les matrices $H_\varphi^{(n)}$ définissent un produit scalaire, ce qui permet alors d'écrire leurs coefficients comme les moments d'une variable aléatoire discrète.

Les trois parties du problème sont relativement indépendantes les unes des autres, mais un même exemple apparaît à la fin de chacune d'entre elles.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Partie 1 : premiers exemples

1. Soit $H = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 3 & -4 \\ 3 & -4 & 5 \end{pmatrix}$.

On note h l'endomorphisme de $\mathbb{R}^3$ dont H est la matrice dans la base canonique.

a) L'espace $\mathbb{R}^3$ est muni de son produit scalaire usuel, noté $\langle \cdot, \cdot \rangle$, pour lequel la base canonique est orthonormale.

i) Soit a un vecteur non nul de $\mathbb{R}^3$ et D la droite engendrée par a . Déterminer, pour tout vecteur x de $\mathbb{R}^3$, la projection orthogonale de x sur la droite D .

ii) Énoncer le théorème spectral. En déduire que le noyau et l'image de h sont deux sous-espaces supplémentaires et orthogonaux de $\mathbb{R}^3$.

iii) Trouver la matrice dans la base canonique de $\mathbb{R}^3$ du projecteur orthogonal sur l'image de h .

b) i) Calculer la trace de l'endomorphisme $h \circ h$.

ii) En déduire le spectre de l'endomorphisme h .

2. Dans cette question, n désigne un nombre entier supérieur ou égal à 3 et on note φ la fonction définie sur $\mathbb{N}$ par :

$$\forall k \in \mathbb{N}, \varphi(k) = (-1)^k (k+1).$$

On observera que $H_\varphi^{(3)}$ est la matrice de la question précédente.

a) Soit τ la fonction définie sur $\mathbb{N}$ par :

$$\forall k \in \mathbb{N}, \tau(k) = k+1.$$

Écrire les matrices $H_\varphi^{(n)}$ et $H_\tau^{(n)}$ et démontrer qu'elles sont semblables.

b) Soit $\mathcal{B}_n = (e_1, e_2, \dots, e_n)$ la base canonique de $\mathbb{R}^n$.

On note g l'endomorphisme de $\mathbb{R}^n$ dont la matrice dans la base $\mathcal{B}_n$ est $H_\tau^{(n)}$.

i) Montrer que l'image de g est le sous-espace vectoriel F de $\mathbb{R}^n$ engendré par les deux vecteurs :

$$\begin{cases} f_1 = \sum_{k=1}^n e_k \\ f_2 = \sum_{k=1}^n k e_k \end{cases}.$$

ii) Justifier la stabilité de F par g et démontrer que la matrice dans la base (f_1, f_2) de l'endomorphisme de F induit par g est :

$$G = \begin{pmatrix} \frac{n(n-1)}{2} & \frac{n(n^2-1)}{3} \\ n & \frac{n(n+1)}{2} \end{pmatrix}.$$

c) Déduire des résultats précédents le spectre de $H_\varphi^{(n)}$.

3. Dans cette question, n désigne un nombre entier strictement positif et on note φ la fonction définie sur $\mathbb{N}$ par :

$$\forall k \in \mathbb{N}, \varphi(k) = k!.$$

a) Soit $C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ une matrice-colonne de $\mathcal{C}_n$.

i) Justifier la convergence de l'intégrale $\int_0^{+\infty} (\sum_{i=1}^n c_i t^{i-1})^2 e^{-t} dt$.

ii) Justifier l'égalité : $\int_0^{+\infty} (\sum_{i=1}^n c_i t^{i-1})^2 e^{-t} dt = {}^t C H_\varphi^{(n)} C$.

b) Dédurre de ce qui précède que toutes les valeurs propres de $H_\varphi^{(n)}$ sont strictement positives.

4. Par un procédé similaire à celui utilisé dans la question précédente, démontrer que toutes les valeurs propres de la matrice $H_\varphi^{(n)}$ sont strictement positives dans le cas où la fonction φ est définie par :

$$\forall k \in \mathbb{N}, \varphi(k) = \frac{1}{k+1}.$$

Partie 2 : les formes bilinéaires Δ_n

Tous les polynômes considérés dans la suite du problème sont à coefficients réels et, pour tout entier naturel p , on note E_p l'espace vectoriel des polynômes de degré au plus p (E_p est donc une notation simplifiée pour $\mathbb{R}_p[X]$).

On considère une application $\varphi : \mathbb{N} \rightarrow \mathbb{R}$, à laquelle on associe pour tout entier naturel n les applications Δ_n et δ_n définies sur $E_n \times E_n$ et E_{2n} de la manière suivante.

- Pour tout polynôme $A = \sum_{i=0}^n a_i X^i \in E_n$ et tout polynôme $B = \sum_{j=0}^n b_j X^j \in E_n$, on pose :

$$\Delta_n(A, B) = \sum_{i=0}^n \sum_{j=0}^n a_i b_j \varphi(i+j).$$

- Pour tout polynôme $Q = \sum_{k=0}^{2n} c_k X^k \in E_{2n}$, on pose : $\delta_n(Q) = \sum_{k=0}^{2n} c_k \varphi(k)$.

1. Soit $n \in \mathbb{N}^*$.

a) Préciser la dimension du $\mathbb{R}$ -espace vectoriel des formes bilinéaires symétriques sur $E_n \times E_n$.

b) Vérifier que l'application Δ_n est une forme bilinéaire symétrique sur $E_n \times E_n$.

c) Établir, pour tout élément (A, B) de $E_n \times E_n$, l'égalité :

$$\Delta_n(A, B) = \delta_n(AB).$$

d) Quelle est la dimension de l'espace vectoriel des formes bilinéaires symétriques sur $E_n \times E_n$ qui peuvent s'écrire $(A, B) \mapsto \delta(AB)$, avec δ forme linéaire sur E_{2n} ?

2. Dans cette question, on considère un nombre entier naturel d , au moins égal à deux, et une variable aléatoire discrète Y , définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, prenant d valeurs distinctes $y_1, y_2, \dots, y_d$ avec les probabilités respectives $p_1, p_2, \dots, p_d$, strictement positives et de somme égale à 1.

On se place alors dans le cas où la fonction φ est définie par :

$$\forall k \in \mathbb{N}, \varphi(k) = \mathbf{E}(Y^k),$$

où $\mathbf{E}(Y^k)$ désigne l'espérance de la variable aléatoire Y^k .

- a) Pour tout polynôme Q de E_{2n} , vérifier l'égalité : $\delta_n(Q) = \mathbf{E}(Q(Y))$.
- b) En déduire une condition nécessaire et suffisante, portant sur n et d , pour que la forme bilinéaire Δ_n soit un produit scalaire sur $E_n \times E_n$.
3. Dans cette question, on suppose que la fonction φ est définie par :

$$\forall k \in \mathbb{N}, \varphi(k) = \frac{1}{k+1}.$$

- a) i) Pour tous (a_0, a_1, a_2) et (b_0, b_1, b_2) éléments de $\mathbb{R}^3$, vérifier l'égalité :

$$\Delta_2(a_2 X^2 + a_1 X + a_0, b_2 X^2 + b_1 X + b_0) = {}^t C_A H_\varphi^{(3)} C_B$$

$$\text{où } C_A = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \text{ et } C_B = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}.$$

- ii) En déduire que Δ_2 est un produit scalaire.

b) À l'aide du procédé d'orthonormalisation de Schmidt, construire à partir de la base canonique $\mathcal{B}_2 = (1, X, X^2)$ de E_2 une base orthonormale $\mathcal{B}'_2$ de E_2 pour le produit scalaire Δ_2 .

c) On note M la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}'_2$ et N la matrice de passage de $\mathcal{B}'_2$ à $\mathcal{B}_2$.

- i) Montrer que la matrice N est triangulaire.

ii) Calculer la matrice ${}^t M H_\varphi^{(3)} M$ et en déduire l'égalité : $H_\varphi^{(3)} = {}^t N N$.

d) Généraliser la démarche précédente pour prouver que, pour tout entier n strictement positif, il existe une matrice triangulaire $T^{(n)}$ de $\mathcal{M}_n$ telle que :

$$H_\varphi^{(n)} = {}^t T^{(n)} T^{(n)}.$$

Partie 3 : polynômes positifs et matrices de moments

Dans cette partie, on utilise les mêmes notations que dans la partie 2.

On considère donc encore une application $\varphi : \mathbb{N} \rightarrow \mathbb{R}$, à laquelle on associe pour tout n des applications Δ_n et δ_n définies comme précédemment.

On dit qu'un polynôme P est *positif* s'il est distinct du polynôme nul et si $P(x)$ est positif ou nul pour tout réel x .

1.
 - a) Montrer que l'ordre de multiplicité de toute racine réelle α d'un polynôme positif P est pair.
 - b) Montrer que tout polynôme positif Q de degré 2 est somme de deux carrés de polynômes, c'est-à-dire qu'il existe un couple (A, B) de polynômes tels que $Q = A^2 + B^2$.
 - c) Pour des polynômes A, B, C, D , écrire le polynôme $(AC + BD)^2 + (AD - BC)^2$ comme produit de deux sommes de deux carrés de polynômes.
 - d) En utilisant sa décomposition en produit de facteurs irréductibles de $\mathbb{R}[X]$, déduire des résultats précédents que tout polynôme positif est somme de deux carrés de polynômes.
 - e) Soit $n \in \mathbb{N}$. Montrer que Δ_n est un produit scalaire sur $E_n \times E_n$ si et seulement si, pour tout polynôme positif P de E_{2n} , on a : $\delta_n(P) > 0$.

Jusqu'à la fin du problème, on suppose que la fonction φ vérifie l'égalité $\varphi(0) = 1$.

On considère un entier n supérieur ou égal à 2, et on suppose que l'application Δ_n associée à φ est un produit scalaire sur $E_n \times E_n$.

On note $(P_0, P_1, \dots, P_n)$ la base orthonormale de E_n (pour ce produit scalaire) obtenue par le procédé d'orthonormalisation de Schmidt à partir de $(1, X, \dots, X^n)$.

2.
 - a) Justifier l'orthogonalité de P_n avec tous les polynômes de degré inférieur ou égal à $n - 2$.
 - b) En déduire que P_n est scindé à racines simples. *On pourra raisonner par l'absurde.*
3. On note $\alpha_1, \dots, \alpha_n$ les n racines de P_n , réelles et distinctes deux à deux, d'après ce qui précède.

Pour tout entier k compris entre 1 et n , on note : $L_k = \prod_{\substack{i=1 \\ i \neq k}}^n \frac{X - \alpha_i}{\alpha_k - \alpha_i}$.

a) Montrer que $(L_1, L_2, \dots, L_n)$ est une base de E_{n-1} et calculer $\sum_{i=1}^n L_i$.

b) Soit $Q \in E_{2n-1}$.

i) Justifier l'existence d'un couple $(A, R) \in E_{n-1} \times E_{n-1}$ tel que : $Q = P_n A + R$.

ii) Vérifier que : $\delta_n(Q) = \sum_{i=1}^n Q(\alpha_i) \delta_n(L_i)$.

c) Pour tout entier k compris entre 1 et n , on pose : $p_k = \delta_n(L_k)$.

Démontrer que $p_1, p_2, \dots, p_n$ sont strictement positifs et de somme égale à 1.

4.
 - a) Déduire de ce qui précède qu'il existe une variable aléatoire discrète Z vérifiant la propriété :

$$\forall k \in \llbracket 1, 2n - 1 \rrbracket, \quad \mathbf{E}(Z^k) = \varphi(k) .$$

b) On revient ici au cas où la fonction φ est définie par : $\forall k \in \mathbb{N}, \varphi(k) = \frac{1}{k+1}$.

Déterminer le nombre minimal de valeurs prises par une variable aléatoire discrète Z vérifiant la propriété précédente.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2015

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet se compose d'un problème et d'un exercice pouvant être traités de manière indépendante. Les différentes parties du problème peuvent également, dans une certaine mesure, être traitées de manière indépendante, quitte à admettre les résultats des parties précédentes. Une liste de résultats susceptibles d'être utiles, et pouvant être utilisés sans qu'il soit nécessaire de les redémontrer, est donnée au début.

La qualité de la rédaction et la précision des réponses apportées entrent pour une part importante dans la notation. Les réponses absurdes sont pénalisées.

Rappel de résultats

Ci-après figurent plusieurs résultats que l'on pourra utiliser sans qu'il soit nécessaire de les redémontrer. Par souci de simplicité, certains énoncés sont donnés dans des cas particuliers (variables aléatoires ne prenant qu'un nombre fini de valeurs, ou variables aléatoires à valeurs entières), et non pas dans le cadre le plus général possible.

1. **Inégalité triangulaire pour l'espérance** : pour toute variable aléatoire réelle X ne prenant qu'un nombre fini de valeurs, on a l'inégalité :

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|).$$

2. **Loi faible des grands nombres** : étant donnée une suite de variables aléatoires réelles $(X_n)_{n \geq 0}$ ne prenant qu'un nombre fini de valeurs, on a, sous l'hypothèse que les variables sont indépendantes et de même loi, la loi faible des grands nombres : pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left| \frac{X_0 + \dots + X_{n-1}}{n} - \mathbb{E}(X) \right| \geq \epsilon \right) = 0.$$

3. **Espérance d'une fonction d'une variable aléatoire** : étant donnée une variable aléatoire réelle X ne prenant qu'un nombre fini de valeurs distinctes $\{x_1, \dots, x_m\}$, et une fonction $h : \mathbb{R} \rightarrow \mathbb{R}$, on a l'identité :

$$\mathbb{E}(h(X)) = \sum_{q=1}^m h(x_q) \cdot \mathbb{P}(X = x_q).$$

4. **Loi d'une somme de deux variables aléatoires indépendantes** : étant données deux variables aléatoires réelles X et Y à valeurs dans $\mathbb{N}$, on a, pour tout $k \in \mathbb{N}$, l'identité :

$$\mathbb{P}(X + Y = k) = \sum_{j=0}^k \mathbb{P}(X = j) \cdot \mathbb{P}(Y = k - j).$$

5. **Loi uniforme sur $[0, 1]$** : la loi uniforme sur l'intervalle $[0, 1]$ est la loi associée à la densité f définie par

$$f(x) = \begin{cases} 1 & \text{si } x \in [0, 1] \\ 0 & \text{si } x \notin [0, 1] \end{cases}.$$

Pour deux nombres a, b tels que $0 \leq a \leq b \leq 1$, on a alors l'identité

$$\mathbb{P}(a < U \leq b) = b - a.$$

6. **Borne sur la probabilité d'une réunion d'événements** : étant donné un nombre $n \geq 1$ d'événements $A_1, \dots, A_n$ définis sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, on a l'inégalité :

$$\mathbb{P} \left(\bigcup_{i=1}^n A_i \right) \leq \sum_{i=1}^n \mathbb{P}(A_i).$$

7. **Développement de Taylor de l'exponentielle** : au voisinage de $x = 0$, on a le développement limité suivant :

$$e^x = 1 + x + \frac{x^2}{2} + o(x^2).$$

Problème

Dans ce problème, on étudie, sur un modèle mathématique très simplifié, les conséquences à long terme de la répartition d'un investissement entre un placement sans risque, garantissant un taux de croissance constant, et un placement risqué, dont le taux de croissance est modélisé par une variable aléatoire.

On considère un pas de temps $\delta > 0$, et une succession de dates $t_0, t_1, \dots$ définies par $t_n = n\delta$ pour tout $n \geq 0$. On suppose que, entre deux dates successives, le placement sans risque voit son montant systématiquement multiplié par une constante $\rho \geq 1$. Autrement dit, pour tout $n \geq 0$, x euros investis à la date t_n dans le placement sans risque se transforment en $x \cdot \rho$ euros à la date t_{n+1} . Pour ce qui est du placement risqué, son montant est multiplié entre deux dates successives par une variable aléatoire : pour tout $n \geq 0$, x euros investis à la date t_n dans le placement risqué se transforment en $x \cdot X_n$ euros à la date t_{n+1} , où X_n est une variable aléatoire. On fait l'hypothèse que les variables $(X_n)_{n \geq 0}$ sont indépendantes et de même loi, et ne peuvent prendre que deux valeurs distinctes h (valeur « haute ») et b (valeur « basse »), leur loi commune étant définie par

$$\begin{cases} \mathbb{P}(X_n = h) = p \\ \mathbb{P}(X_n = b) = q \end{cases},$$

où :

- h et b sont deux nombres réels strictement positifs vérifiant la relation $b < h$.
- p et q sont deux nombres réels strictement positifs vérifiant la relation $p + q = 1$.

La stratégie d'investissement que l'on étudie consiste, partant d'un capital initial non-aléatoire $C_0 > 0$ à la date t_0 , à répartir à chaque date t_n le montant disponible en une proportion α investie dans le placement risqué, et une proportion $1 - \alpha$ dans le placement sans risque, le nombre α étant une constante fixée dans $[0, 1]$. On suppose en outre qu'il n'y a au cours du temps ni apport extérieur (hormis le capital initial C_0) ni retrait. Dans la suite, on note C_n le montant disponible à la date C_n obtenu en suivant la stratégie d'investissement étudiée.

Partie I – Quelques résultats généraux

Question 1) Exprimer, pour tout $n \geq 0$, la valeur de C_{n+1} en fonction de C_n , X_n , α et ρ .

Question 2) En déduire la relation, valable pour tout $n \geq 1$:

$$C_n = C_0 \cdot \prod_{k=0}^{n-1} [\alpha X_k + (1 - \alpha)\rho].$$

Question 3) On définit la fonction $\phi : [0, 1] \rightarrow \mathbb{R}$ par

$$\phi(\alpha) = \mathbb{E}(\alpha X_0 + (1 - \alpha)\rho). \quad (1)$$

Déduire de ce qui précède la relation, valable pour tout $n \geq 1$:

$$\mathbb{E}(C_n) = C_0 \times (\phi(\alpha))^n.$$

Question 4) Montrer que, pour tout $n \geq 0$, on a $C_n > 0$, puis établir pour tout $n \geq 1$ la relation :

$$\ln C_n = \ln C_0 + \sum_{k=0}^{n-1} \ln(\alpha X_k + (1 - \alpha)\rho).$$

Question 5) On définit la fonction $g : [0, 1] \rightarrow \mathbb{R}$ par

$$g(\alpha) = \mathbb{E}(\ln(\alpha X_0 + (1 - \alpha)\rho)). \quad (2)$$

Montrer que l'on a la propriété suivante : pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left|\frac{\ln C_n}{n} - g(\alpha)\right| \geq \epsilon\right) = 0.$$

Indication : Exploiter le résultat de la question 4) en faisant appel à la loi des grands nombres.

Question 6) En déduire que l'on peut écrire, pour $n \geq 1$,

$$C_n = e^{n(g(\alpha) + \epsilon_n)},$$

où $(\epsilon_n)_{n \geq 1}$ est une suite de variables aléatoires vérifiant la propriété suivante : pour tout $\epsilon > 0$
 $\lim_{n \rightarrow +\infty} \mathbb{P}(|\epsilon_n| \geq \epsilon) = 0$.

Partie II – Taux de croissance optimal

On introduit la notation $\mu = \mathbb{E}(X_0)$, et, dans toute cette partie, on fait les hypothèses suivantes :

- $b < \rho < h$
- $\mu > \rho$

On introduit également la notation

$$\alpha^* = \frac{\rho(\mu - \rho)}{(h - \rho)(\rho - b)}.$$

Question 7) Donner une expression explicite de $g(\alpha)$ en fonction de α et des autres paramètres du problème (ρ, h, b, p, q) .

Question 8) Montrer que l'on a toujours $g(0) \geq 0$.

Question 9) Montrer que, si $\alpha^* < 1$, la fonction g est strictement croissante sur $[0, \alpha^*]$, puis strictement décroissante sur $[\alpha^*, 1]$, et possède donc un unique maximum sur $[0, 1]$ atteint en α^* .

Question 10) Montrer que, si $\alpha^* \geq 1$, la fonction g est strictement croissante sur $[0, 1]$, et possède donc un unique maximum sur $[0, 1]$ atteint en 1.

Question 11) Quelle valeur de α doit-on choisir (selon vous) afin de maximiser la croissance à long terme du montant disponible ? (Justifier la réponse.)

Partie III – Un exemple

On suppose que les paramètres prennent les valeurs suivantes : $h = 1,4$; $b = 0,7$; $\rho = 1,04$; $p = q = 0,5$. On donne la valeur approchée (à $\pm 10^{-3}$ près) $\alpha^* \approx 0,085$.

Question 12) Montrer que sur cet exemple on a $\phi(1) > 1$ et $g(1) < 0$.

Question 13) En déduire que, avec le choix $\alpha = 1$, on a les deux propriétés suivantes :

$$\lim_{n \rightarrow +\infty} \mathbb{E}(C_n) = +\infty,$$

et, pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(C_n \geq \epsilon) = 0.$$

Comment expliquer l'occurrence simultanée de ces deux propriétés apparemment contradictoires ?

Partie IV – Probabilité de franchissement d'un seuil

On suppose dans cette partie que $h = e^\ell$ et $b = e^{-\ell}$, où ℓ est un nombre réel strictement positif, et l'on suppose également que $p > q$. On définit $Z_0 = 0$ et, pour tout $n \geq 1$,

$$Z_n = \sum_{k=0}^{n-1} \ln X_k.$$

Pour tout entier $m \geq 0$, on introduit l'événement A_m défini par

$$A_m = \{\text{il existe } n \geq 0 \text{ tel que } Z_n = -m \times \ell\}$$

et la probabilité

$$u_m = \mathbb{P}(A_m).$$

Question 14) Montrer que $u_0 = 1$.

Question 15) Montrer que, pour tout $m \geq 1$, on a les identités :

$$\mathbb{P}(A_m | X_0 = h) = \mathbb{P}(A_{m+1}) \text{ et } \mathbb{P}(A_m | X_0 = b) = \mathbb{P}(A_{m-1}).$$

Question 16) En déduire la relation, valable pour tout $m \geq 1$:

$$u_m = pu_{m+1} + qu_{m-1}.$$

Question 17) En déduire qu'il existe des constantes c_1 et c_2 telles que, pour tout $m \geq 0$,

$$u_m = c_1 + c_2 \left(\frac{q}{p}\right)^m.$$

On admettra dans la suite que la propriété suivante est vérifiée :

$$\lim_{m \rightarrow +\infty} u_m = 0. \quad (3)$$

Question 18) Montrer que, pour tout $m \geq 1$,

$$u_m = \left(\frac{q}{p}\right)^m.$$

Indication : utiliser (3) et les résultats des questions 14) et 17).

Partie V – Passage à la limite

Comme dans la partie IV, on suppose que $h = e^\ell$ et $b = e^{-\ell}$. On suppose en outre que les paramètres du modèle sont reliés au pas de temps δ de la manière suivante :

$$\ell = \sigma \cdot \sqrt{\delta}, \quad p = \frac{1}{2} \left(1 + \frac{\left(m - \frac{\sigma^2}{2}\right)}{\sigma} \sqrt{\delta} \right), \quad q = \frac{1}{2} \left(1 - \frac{\left(m - \frac{\sigma^2}{2}\right)}{\sigma} \sqrt{\delta} \right), \quad \rho = 1 + r\delta,$$

où r , m et σ sont trois constantes vérifiant $m > r \geq 0$ et $\sigma > 0$, et δ est supposé suffisamment petit pour que p et q soient strictement positifs.

Question 19) Montrer que l'on a les propriétés suivantes lorsque $\delta \rightarrow 0^+$:

$$(i) \quad \mathbb{E}(X_0) = 1 + m\delta + o(\delta)$$

(ii) $\mathbb{V}(X_0) = \sigma^2 \delta + o(\delta)$ (la notation $\mathbb{V}$ désigne la variance)

(iii) $\lim_{\delta \rightarrow 0^+} \alpha^* = \frac{m-r}{\sigma^2}$.

Indication : Expliciter les valeurs de $\mathbb{E}(X_0)$, $\mathbb{V}(X_0)$ et α^* , puis effectuer un développement limité en δ .

Question 20) On suppose pour simplifier que $r = 0$ et que $\frac{m-r}{\sigma^2} = 1$, de telle sorte que l'on a $\lim_{\delta \rightarrow 0^+} \alpha^* = 1$. En supposant que l'on choisit exactement $\alpha = 1$, montrer que l'on a, pour tout $\theta \in]0, 1[$, le résultat suivant :

$$\lim_{\delta \rightarrow 0^+} \mathbb{P} \left[\text{il existe } n \geq 0 \text{ tel que } \left(\frac{C_n}{C_0} \right) \leq \theta \right] = \theta. \quad (4)$$

Indication : Relier la probabilité figurant dans (4) à la probabilité étudiée dans la partie IV.

Question 21) On peut montrer dans un cadre très général qu'une relation analogue à (4) est valable, lorsque l'on choisit d'investir à chaque pas de temps une proportion α^* du montant disponible dans le placement risqué. Commenter la relation (4) du point de vue du risque associé à une telle stratégie d'investissement à long terme.

Exercice

On considère un entier $N \geq 1$, et N variables aléatoires $B_1, \dots, B_N$ indépendantes. On suppose en outre que, pour tout $i \in \{1, \dots, N\}$, la variable aléatoire B_i suit une loi de Bernoulli de paramètre p_i , où $p_i \in [0, 1]$; autrement dit, la loi de B_i est caractérisée par :

$$\begin{cases} \mathbb{P}(B_i = 1) = p_i \\ \mathbb{P}(B_i = 0) = 1 - p_i \end{cases}.$$

On introduit ensuite la variable S_N définie par

$$S_N = \sum_{i=1}^N B_i,$$

et l'on considère également une variable aléatoire T suivant la loi de Poisson de paramètre λ , où λ est défini par

$$\lambda = \sum_{i=1}^N p_i.$$

L'objectif de cet exercice est d'obtenir une borne sur l'écart entre la loi de S_N et la loi de T .

Question A) On considère deux variables aléatoires indépendantes X et Y , X suivant la loi de Poisson de paramètre β , et Y la loi de Poisson de paramètre γ . Montrer que la loi de la variable aléatoire $X + Y$ est la loi de Poisson de paramètre $\beta + \gamma$.

Indication : Calculer la valeur de $\mathbb{P}(X + Y = k)$ pour tout entier $k \geq 0$.

Dans la suite de l'exercice, on considère N variables aléatoires $U_1, \dots, U_N$ indépendantes et de même loi uniforme sur $[0, 1]$.

Question B) Pour tout $i \in \{1, \dots, N\}$, on définit la variable aléatoire C_i par

$$C_i = \begin{cases} 1 & \text{si } U_i > 1 - p_i \\ 0 & \text{si } U_i \leq 1 - p_i \end{cases}.$$

Montrer que, pour tout $i \in \{1, \dots, N\}$, la variable aléatoire C_i suit la loi de Bernoulli de paramètre p_i .

Indication : calculer la probabilité $\mathbb{P}(C_i = 1)$.

Question C) Pour tout $k \in \mathbb{N}$, et tout $p \in [0, 1]$, on note $g_p(k) = \sum_{j=0}^k e^{-p} \frac{p^j}{j!}$, et l'on pose $g_p(-1) = 0$. Pour tout $i \in \{1, \dots, N\}$, on définit ensuite la variable aléatoire D_i en stipulant que, pour tout $k \in \mathbb{N}$,

$$D_i = k \text{ si et seulement si } g_{p_i}(k-1) < U_i \leq g_{p_i}(k).$$

Montrer que, pour tout $i \in \{1, \dots, N\}$, la variable aléatoire D_i suit la loi de Poisson de paramètre p_i .

Indication : calculer pour tout k la probabilité $\mathbb{P}(D_i = k)$.

Question D) Représenter sur une figure l'ensemble des valeurs de la variable U_i pour lesquelles on a l'égalité $C_i = D_i$, en tenant compte de l'inégalité $1 - p_i \leq e^{-p_i}$.

Question E) En utilisant l'inégalité $1 - p_i \leq e^{-p_i}$, montrer que, pour tout $i \in \{1, \dots, N\}$, on a l'inégalité

$$\mathbb{P}(C_i = D_i) \geq 1 - p_i^2. \quad (5)$$

Question F) On introduit l'événement $A = \{\forall i \in \{1, \dots, N\}, C_i = D_i\}$. Montrer que

$$\mathbb{P}(\bar{A}) \leq \sum_{i=1}^N p_i^2, \quad (6)$$

où $\bar{A}$ désigne l'événement complémentaire de A .

Dans la suite de l'exercice, on considère les variables aléatoires

$$C = \sum_{i=1}^N C_i \text{ et } D = \sum_{i=1}^N D_i.$$

Question G) Montrer que la loi de C est identique à la loi de S_N , et que la loi de D est identique à la loi de T .

Question H) On considère un sous-ensemble $H \subset \mathbb{N}$, et l'on introduit les variables aléatoires V et W définies respectivement par

$$V = \begin{cases} 1 & \text{si } C \in H \\ 0 & \text{si } C \notin H \end{cases} \text{ et } W = \begin{cases} 1 & \text{si } D \in H \\ 0 & \text{si } D \notin H \end{cases}.$$

Montrer que l'on a les identités

$$\mathbb{P}(C \in H) = \mathbb{E}(V) \text{ et } \mathbb{P}(D \in H) = \mathbb{E}(W).$$

Question I) Dédurre de la question précédente que l'on a l'inégalité

$$|\mathbb{P}(C \in H) - \mathbb{P}(D \in H)| \leq \mathbb{E}(|V - W|).$$

Question J) On introduit la variable aléatoire Y définie par

$$Y = \begin{cases} 1 & \text{si } A \text{ n'est pas réalisé} \\ 0 & \text{si } A \text{ est réalisé} \end{cases}.$$

Montrer que l'on a l'inégalité $|V - W| \leq Y$.

Question K) En faisant la synthèse des questions précédentes, montrer que

$$|\mathbb{P}(S_N \in H) - \mathbb{P}(T \in H)| \leq \sum_{i=1}^N p_i^2. \quad (7)$$

Expliquer en quoi ce résultat répond à l'objectif de l'exercice présenté au début.

Question L) On se donne un nombre réel positif h tel que $h \leq N$, et l'on fait l'hypothèse que $p_i = h/N$ pour tout $i \in \{1, \dots, N\}$. Quelles sont alors les lois respectives de S_N et T ? Quelle est dans ce cas la valeur du membre de droite de l'inégalité (7)? Comment se comporte celui-ci lorsque $N \rightarrow \infty$?

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2016

Épreuve de mathématiques

Durée : 4h

L'objet du problème est l'étude, dans différents contextes, de la convergence et, parfois, de la valeur de la somme d'une série de terme général $\frac{\varepsilon_n}{n}$ où, pour tout entier naturel n non nul, ε_n vaut 1 ou -1 .

Partie I : Quelques cas déterministes

On note, pour tout entier naturel n non nul, $H_n = \sum_{k=1}^n \frac{1}{k}$.

1. On pose, pour tout entier naturel n non nul, $\gamma_n = H_n - \ln n$.
Montrer que la suite de terme général $(\gamma_n)_{n \in \mathbb{N}^*}$ est convergente. On note γ sa limite.
2. a) Établir, pour tout entier naturel N non nul, l'égalité : $\sum_{n=1}^{2N} \frac{(-1)^n}{n} = H_N - H_{2N}$.
b) En déduire la convergence de la série $\sum \frac{(-1)^n}{n}$ et la valeur de la somme $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n}$.
3. On suppose dans cette question, que, pour tout entier naturel n , $\varepsilon_{3n+1} = \varepsilon_{3n+2} = 1$ et $\varepsilon_{3n+3} = -1$. Quelle est la nature de la série $\sum \frac{\varepsilon_n}{n}$?
4. On suppose dans cette question, que, pour tout entier naturel n ,

$$\varepsilon_{4n+1} = \varepsilon_{4n+2} = 1 \quad \text{et} \quad \varepsilon_{4n+3} = \varepsilon_{4n+4} = -1.$$

- a) Prouver, pour tout entier naturel N , l'égalité :

$$\sum_{n=0}^N \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) = \int_0^1 \frac{1-x^{4N+4}}{1+x^2} dx.$$

- b) En déduire l'égalité : $\sum_{n=0}^{+\infty} \left(\frac{1}{4n+1} - \frac{1}{4n+3} \right) = \frac{\pi}{4}$.

- c) Calculer, de même, la somme $\sum_{n=0}^{+\infty} \left(\frac{1}{4n+2} - \frac{1}{4n+4} \right)$ et en déduire que la série $\sum \frac{\varepsilon_n}{n}$ est convergente avec $\sum_{n=1}^{+\infty} \frac{\varepsilon_n}{n} = \frac{\pi}{4} + \frac{1}{2} \ln 2$.

Dans toute la suite de cette partie, pour tout entier naturel p non nul, on considère la suite $(\varepsilon_n)_{n \in \mathbb{N}^*}$ telle que, pour tout entier naturel n ,

$$\varepsilon_{2pn+1} = \varepsilon_{2pn+2} = \dots = \varepsilon_{2pn+p} = 1 \quad \text{et} \quad \varepsilon_{2pn+p+1} = \varepsilon_{2pn+p+2} = \dots = \varepsilon_{2pn+2p} = -1.$$

La suite $(\varepsilon_n)_{n \in \mathbb{N}^*}$ est donc constituée de p termes égaux à 1, suivis de p termes égaux à -1, suivis de p termes égaux à 1, etc.

On pose, pour tout entier naturel n non nul, $S_n = \sum_{k=1}^n \frac{\varepsilon_k}{k}$.

5. a) Prouver, pour tout entier naturel n , la majoration : $\sum_{k=2pn+1}^{2pn+2p} \frac{\varepsilon_k}{k} \leq \frac{1}{2n^2}$.

- b) En déduire la convergence de la suite $(S_{2pn})_{n \in \mathbb{N}^*}$.

- c) Prouver la convergence de la série $\sum \frac{\varepsilon_n}{n}$.

On note Σ_p la somme de cette série c'est-à-dire $\Sigma_p = \sum_{n=1}^{+\infty} \frac{\varepsilon_n}{n}$.

6. a) Établir, pour tout entier naturel N non nul, l'égalité :

$$\sum_{n=0}^N \left(\frac{1}{2pn+1} + \dots + \frac{1}{2pn+p} - \frac{1}{2pn+p+1} - \dots - \frac{1}{2pn+p+p} \right) = \sum_{n=0}^N \int_0^1 (1+x+\dots+x^{p-1})(1-x^p)x^{2pn} dx.$$

b) En déduire l'égalité : $\lim_{N \rightarrow +\infty} S_{2p(N+1)} = \int_0^1 \frac{1+x+\dots+x^{p-1}}{1+x^p} dx$.

c) Conclure à l'égalité : $\Sigma_p = \int_0^1 \frac{1+x+\dots+x^{p-1}}{1+x^p} dx$.

7. Calcul de Σ_3

a) Soit $(a, b) \in \mathbb{R} \times \mathbb{R}^*$. Vérifier que la fonction $x \mapsto \frac{1}{b} \arctan\left(\frac{x-a}{b}\right)$ est une primitive sur $\mathbb{R}$ de la fonction $x \mapsto \frac{1}{(x-a)^2 + b^2}$.

b) En déduire la valeur de $\int_0^1 \frac{1}{x^2 - x + 1} dx$ puis conclure à l'égalité : $\Sigma_3 = \frac{2\pi}{3\sqrt{3}} + \frac{1}{3} \ln 2$.

8. Une relation de récurrence

a) Justifier la convergence de l'intégrale $\int_0^{+\infty} \frac{1+x+\dots+x^{p-2}}{1+x^p} dx$ et établir l'égalité :

$$(1) \quad \Sigma_p = \frac{\ln 2}{p} + \frac{1}{2} \int_0^{+\infty} \frac{1+x+\dots+x^{p-2}}{1+x^p} dx.$$

b) Établir l'égalité : $\int_0^{+\infty} \frac{x(1+x^2+x^4+\dots+x^{2p-4})}{1+x^{2p}} dx = \frac{1}{2} \int_0^{+\infty} \frac{(1+u+u^2+\dots+u^{p-2})}{1+u^p} du$.

c) En déduire l'égalité : $\Sigma_{2p} = \frac{1}{2} \int_0^{+\infty} \frac{1+x^2+x^4+\dots+x^{2p-2}}{1+x^{2p}} dx + \frac{1}{2} \Sigma_p$.

9. Calcul de $\int_0^{+\infty} \frac{x^{2q}}{1+x^{2p}} dx$

a) Soit A un réel strictement positif et f est une fonction continue définie sur $\mathbb{R}$ à valeurs dans $\mathbb{C}$ dont on note f_1 (resp. f_2) la partie réelle (resp. imaginaire).

On rappelle (ou on admet) que $\int_{-A}^A f(t) dt$ désigne la somme $\int_{-A}^A f_1(t) dt + i \int_{-A}^A f_2(t) dt$.

Soit $\alpha \in]0, 2\pi[\setminus \{\pi\}$. Établir l'égalité :

$$\lim_{A \rightarrow +\infty} \int_{-A}^A \frac{1}{t - e^{i\alpha}} dt = \begin{cases} i\pi & \text{si } \alpha \in]0, \pi[; \\ -i\pi & \text{sinon.} \end{cases}$$

b) (i) On note, pour tout $k \in \llbracket 0, 2p-1 \rrbracket$, $z_k = \exp\left(\frac{(2k+1)i\pi}{2p}\right)$.

Établir, pour tout entier $r \in \llbracket 0, 2p-1 \rrbracket$, l'existence et l'unicité d'un polynôme de $\mathbb{C}_{2p-1}[X]$, noté L_r , tel que

$$L_r(z_k) = \begin{cases} 0 & \text{si } k \neq r; \\ 1 & \text{sinon.} \end{cases}$$

(ii) Montrer que la famille $(L_0, L_1, \dots, L_{2p-1})$ est une base de $\mathbb{C}_{2p-1}[X]$.

c) Soit $q \in \llbracket 0, p-1 \rrbracket$.

(i) Justifier l'existence de nombres complexes $\lambda_0, \lambda_1, \dots, \lambda_{2p-1}$ tels que, pour tout réel t ,

$$\frac{t^{2q}}{1+t^{2p}} = \sum_{k=0}^{2p-1} \frac{\lambda_k}{t - z_k}.$$

(ii) Établir l'égalité : $\sum_{k=0}^{2p-1} \lambda_k = 0$.

(iii) Prouver, pour tout $k \in \llbracket 0, 2p-1 \rrbracket$, l'égalité : $\lambda_k = -\frac{z_k^{2q+1}}{2p}$.

d) En déduire l'égalité : $\int_0^{+\infty} \frac{x^{2q}}{1+x^{2p}} dx = -\frac{i\pi}{2p} \sum_{k=0}^{p-1} \exp\left(\left[\frac{(2k+1)(2q+1)}{2p}\right] i\pi\right)$.

10. Déterminer la valeur de Σ_4 .

Partie II : Un cas aléatoire

On considère une suite $(Y_n)_{n \in \mathbb{N}^*}$ de variables aléatoires réelles discrètes, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, **indépendantes, centrées** (c'est-à-dire d'espérance nulle) et possédant un moment d'ordre 2.

On **admet** qu'une suite réelle $(u_n)_{n \in \mathbb{N}^*}$ converge si, et seulement si, on a

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N}^* \quad \forall (n, p) \in \mathbb{N}^2 \quad \left(p \geq N \text{ et } n \geq N \Rightarrow |u_p - u_n| \leq \varepsilon \right).$$

On pose, pour tout entier naturel n non nul, $S_n = \sum_{k=1}^n Y_k$ et on note $\mathcal{C}$ l'ensemble des $\omega \in \Omega$ pour lesquels la suite $(S_n(\omega))_{n \in \mathbb{N}^*}$ converge.

A. La convergence presque sûre

1. On pose, pour tout $\varepsilon > 0$, $B(\varepsilon) = \bigcup_{N=1}^{+\infty} \bigcap_{\substack{n \geq N \\ p \geq N}} \left[|S_n - S_p| \leq \varepsilon \right]$.

a) (i) Justifier, pour tout $\varepsilon > 0$, l'appartenance de $B(\varepsilon)$ à $\mathcal{A}$.

(ii) Établir l'égalité : $\mathcal{C} = \bigcap_{\varepsilon > 0} B(\varepsilon)$.

(iii) Comparer les ensembles $B(\varepsilon)$ et $B(\varepsilon')$ quand $0 < \varepsilon < \varepsilon'$.

(iv) Établir l'égalité : $\mathcal{C} = \bigcap_{k=1}^{+\infty} B\left(\frac{1}{k}\right)$ et en déduire que $\mathcal{C} \in \mathcal{A}$.

2. a) Montrer que $\mathbf{P}(\mathcal{C}) = 1$ si, et seulement si, pour tout entier naturel k non nul,

$$\mathbf{P}\left(B\left(\frac{1}{k}\right)\right) = 1.$$

b) En déduire que $\mathbf{P}(\mathcal{C}) = 1$ si, et seulement si, pour tout $\varepsilon > 0$,

$$\mathbf{P}\left(\bigcap_{N=1}^{+\infty} \bigcup_{\substack{n \geq N \\ p \geq N}} \left[|S_p - S_n| > \varepsilon \right]\right) = 0.$$

c) Montrer que $\mathbf{P}(\mathcal{C}) = 1$ si, et seulement si, pour tout $\varepsilon > 0$, $\lim_{N \rightarrow +\infty} \mathbf{P}\left(\bigcup_{\substack{n \geq N \\ p \geq N}} \left[|S_p - S_n| > \varepsilon \right]\right) = 0$.

B. Une inégalité

Quand une variable aléatoire U , définie sur $(\Omega, \mathcal{A}, \mathbf{P})$, a une espérance on note $\mathbf{E}(U)$ sa valeur.

Soit $\varepsilon > 0$ et N un entier naturel non nul. On note T_N l'application qui, à chaque $\omega \in \Omega$, associe l'élément de $\mathbb{N} \cup \{+\infty\}$ défini par

$$T_N(\omega) = \inf \left\{ p \in \mathbb{N}^*; p > N \quad \text{et} \quad |S_p(\omega) - S_N(\omega)| > \varepsilon \right\}$$

(avec la convention $\inf \emptyset = +\infty$).

1. **a)** Soit A un événement. Établir l'égalité : $\mathbf{E}(\mathbf{1}_A) = \mathbf{P}(A)$.
b) Déterminer, pour tout entier naturel N non nul et pour tout entier $p > N$, les valeurs des espérances $\mathbf{E}(S_p - S_N)$ et $\mathbf{E}((S_p - S_N)^2)$ en fonction des moments des Y_k .
2. Exprimer, pour tout entier $k > N$, l'ensemble $[T_N = k]$ à l'aide d'événements liés à différentes variables aléatoires S_i et en déduire que l'application T_N est une variable aléatoire.
3. **a)** Prouver, pour tout entier $k > N$, l'inégalité : $\varepsilon^2 \mathbf{P}([T_N = k]) \leq \mathbf{E}((S_k - S_N)^2 \mathbf{1}_{[T_N = k]})$.
b) Soit p un entier strictement plus grand que N .
Justifier, pour tout $k \in \llbracket N+1, p \rrbracket$, l'indépendance des variables $S_p - S_k$ et $(S_k - S_N) \mathbf{1}_{[T_N = k]}$.
c) En déduire, pour tout $(p, k) \in \mathbb{N}^2$ vérifiant $N < k \leq p$, l'inégalité :

$$\varepsilon^2 \mathbf{P}([T_N = k]) \leq \mathbf{E}((S_p - S_N)^2 \mathbf{1}_{[T_N = k]}).$$

- d)** Prouver, pour tout entier $p > N$, l'inégalité :

$$\varepsilon^2 \sum_{k=N+1}^p \mathbf{P}([T_N = k]) \leq \sum_{i=N+1}^p \mathbf{E}(Y_i^2).$$

4. On suppose, de plus, que la série $\sum \mathbf{E}(Y_m^2)$ converge. Établir l'inégalité :

$$\mathbf{P}\left(\bigcup_{p>N} [|S_p - S_N| > \varepsilon]\right) \leq \frac{1}{\varepsilon^2} \sum_{i=N+1}^{+\infty} \mathbf{E}(Y_i^2).$$

C. Le résultat

On considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires réelles discrètes, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$, **indépendantes** et toutes de **même loi** que X_1 . On suppose que la loi de la variable X_1 est donnée par

$$\mathbf{P}([X_1 = -1]) = \mathbf{P}([X_1 = 1]) = \frac{1}{2}.$$

On pose, pour tout entier naturel n non nul, $S_n = \sum_{k=1}^n \frac{X_k}{k}$.

1. Prouver, pour tout entier naturel N non nul, l'inclusion

$$\bigcup_{\substack{n \geq N \\ p \geq N}} [|S_p - S_n| > \varepsilon] \subset \bigcup_{p > N} [|S_p - S_N| > \frac{\varepsilon}{2}].$$

2. Montrer que, presque sûrement, la série $\sum \frac{X_n}{n}$ converge, c'est-à-dire montrer que l'ensemble des $\omega \in \Omega$ pour lesquels la série $\sum \frac{X_n(\omega)}{n}$ converge est de probabilité 1.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2016

Épreuve à option (A) : Mathématiques

Durée : 4h

Étant donné un entier $n \in \mathbb{N}^*$ et un réel a , on dit qu'une fonction réelle f définie au voisinage de a est *plate à l'ordre n en a* si $f(x) - f(a)$ est négligeable devant $(x - a)^n$ mais pas devant $(x - a)^{n+1}$ lorsque x tend vers a . On dit que f est *ultraplate en a* si $f(x) - f(a)$ est négligeable devant $(x - a)^m$ lorsque x tend vers a , quel que soit l'entier naturel m .

L'objet du problème est l'étude des fonctions plates et de leurs approximations polynomiales. Les quatre parties du sujet sont largement indépendantes, mais les notations de la partie 1 sont utilisées dans les deux parties suivantes.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Partie 1 : étude des espaces $E_n(a)$

On note E le $\mathbb{R}$ -espace vectoriel des applications de classe C^∞ de $\mathbb{R}$ dans $\mathbb{R}$, dont on rappelle qu'il possède aussi une structure d'anneau commutatif et d'algèbre.

Pour tout entier $n \in \mathbb{N}^*$ et tout réel a , on note $E_n(a)$ l'ensemble des fonctions de E telles que la différence $f(x) - f(a)$ soit négligeable devant $(x - a)^n$ lorsque x tend vers a .

1. Soit $f \in E$, $a \in \mathbb{R}$ et $n \in \mathbb{N}^*$.

a) Démontrer, en utilisant une formule de Taylor, que la fonction f appartient à $E_n(a)$ si, et seulement si, pour tout entier k compris entre 1 et n , la dérivée d'ordre k de f en a est nulle :

$$f \in E_n(a) \iff \forall k \in \llbracket 1, n \rrbracket, f^{(k)}(a) = 0.$$

b) Démontrer que la somme $s : x \mapsto \sum_{k=0}^{+\infty} a_k x^k$ d'une série entière de rayon de convergence infini est ultraplate en 0 si, et seulement si, s est une fonction constante.

2. Un exemple de fonction ultraplate en 0

Pour tout $x > 0$, on pose : $b(x) = \exp(-(\ln x)^2)$.

a) Donner l'allure du graphe de b en précisant les coordonnées de ses points d'inflexion.

b) Justifier que b est de classe C^∞ sur $]0, +\infty[$ et prouver que, pour tout $n \in \mathbb{N}^*$, il existe un polynôme $B_n \in \mathbb{R}[X]$ tel que :

$$\forall x > 0, b^{(n)}(x) = \frac{B_n(\ln x)}{x^n} \exp(-(\ln x)^2)$$

(on précisera le degré et le coefficient du monôme dominant de B_n).

c) En déduire que la fonction $c : x \mapsto \begin{cases} b(|x|) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$ est ultraplate en 0. En quels autres points est-elle plate et à quel ordre?

3. Soit $n \in \mathbb{N}^*$.

a) Montrer que, pour tout réel a , $E_n(a)$ est une sous-algèbre de l'algèbre E .

b) L'ensemble $E_n(a)$ est-il un idéal de l'anneau commutatif E ?

4. a) Montrer que, si une fonction f est ultraplate en 0, alors la fonction $x \mapsto f(x-a)$ est ultraplate en a .

b) Construire à l'aide de la fonction c définie plus haut une fonction de E ultraplate à la fois en 0, en +1 et en -1.

Partie 2 : interpolations polynomiales avec ajustement de dérivées

1. a) Trouver, pour tout $n \in \mathbb{N}^*$, l'unique polynôme $P_n \in \mathbb{R}_{n+2}[X]$ vérifiant $P_n(0) = 0$ et $P_n(1) = 1$, et tel que la fonction polynomiale P_n appartienne à $E_n(0) \cap E_1(1)$.

b) Étudier la convergence simple et la convergence uniforme de la suite $(P_n)_{n \in \mathbb{N}^*}$ sur $[0, 1]$.

2. Soit p un nombre entier supérieur ou égal à 2, $a_1, a_2, \dots, a_p$ des nombres réels deux à deux distincts et $n_1, n_2, \dots, n_p$ des nombres entiers strictement positifs.

Soit $m = p - 1 + \sum_{k=1}^p n_k$.

Soit Φ l'application de $\mathbb{R}[X]$ dans $\mathbb{R}^{m+1}$ définie par :

$$\forall P \in \mathbb{R}[X], \quad \Phi(P) = \left(P^{(0)}(a_1), P^{(1)}(a_1), \dots, P^{(n_1)}(a_1), \dots, P^{(0)}(a_p), P^{(1)}(a_p), \dots, P^{(n_p)}(a_p) \right)$$

(où la dérivée 0-ième $P^{(0)}$ de P désigne la fonction polynomiale P elle-même).

a) Vérifier que Φ est linéaire et vérifier que son noyau est un supplémentaire de $\mathbb{R}_m[X]$ dans $\mathbb{R}[X]$.

b) En déduire que pour tout $(\alpha_1, \dots, \alpha_p) \in \mathbb{R}^p$, il existe un unique polynôme P de degré inférieur ou égal à m appartenant à $\bigcap_{k=1}^p E_{n_k}(a_k)$ vérifiant :

$$\forall k \in \llbracket 1, p \rrbracket, P(a_k) = \alpha_k.$$

3. Soit $n \in \mathbb{N}^*$.

a) Justifier l'existence d'une unique fonction polynomiale paire H_n de degré inférieur ou égal à $n + 4$, appartenant à $E_n(0) \cap E_1(-1) \cap E_1(+1)$ et telle que $H_n(0) = 0$, $H_n(-1) = H_n(+1) = 1$.

b) Calculer H_n , selon la parité de n .

c) Étudier la convergence simple et la convergence uniforme de la suite $(H_n)_{n \in \mathbb{N}^*}$ sur l'intervalle ouvert $] -1, +1 [$.

Partie 3 : fonctions génératrices plates en 0

Dans cette partie, on note $(\Omega, \mathcal{A})$ l'espace probabilisable $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

On admet que, quelle que soit la série convergente $\sum p_n$ à termes positifs ou nuls, telle que $\sum_{n=0}^{+\infty} p_n = 1$, il existe une probabilité P sur $(\Omega, \mathcal{A})$ et une variable aléatoire X sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$, à valeurs dans $\mathbb{N}$, vérifiant :

$$\forall n \in \mathbb{N}, P([X = n]) = p_n.$$

1. Dans cette question, G_X désigne la fonction génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N}$, définie sur l'espace probabilisable $(\Omega, \mathcal{A})$ muni d'une probabilité P .
 - a) Soit $n \in \mathbb{N}^*$. Donner une condition nécessaire et suffisante portant sur la fonction de répartition de X pour que G_X soit plate à l'ordre n en 0.
 - b) Démontrer que G_X est ultraplate en 0 si, et seulement si, $P([X = 0]) = 1$.
2. Dans cette question, $\sum a_n x^n$ désigne une série entière de rayon de convergence infini, dont la somme $S(x) = \sum_{n=1}^{+\infty} a_n x^n$ est égale à 1 pour $x = 1$ et dont tous les coefficients a_n sont positifs ou nuls.

Soit P une probabilité sur $(\Omega, \mathcal{A})$ et X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$, à valeurs dans $\mathbb{N}^*$, vérifiant :

$$\forall n \in \mathbb{N}^*, P([X = n]) = a_n.$$

- a) Justifier que X admet un moment à n'importe quel ordre.
 - b) Soit $n \in \mathbb{N}^*$. Montrer que si S est plate à l'ordre n en 0, alors l'espérance de X est supérieure ou égale à $n + 1$. Dans quel cas a-t-on égalité?
3. Soit $n \in \mathbb{N}^*$ et c un réel strictement supérieur à $n + 1$.
 - a) Montrer qu'il existe une variable aléatoire à valeurs dans $\mathbb{N}^*$ dont la fonction génératrice G appartient à $E_n(0)$ et vérifie $G'(1) = c$.
On pourra chercher une variable aléatoire ne prenant que deux valeurs entières distinctes.
 - b) Démontrer que la fonction génératrice G d'une telle variable aléatoire vérifie nécessairement l'inégalité :

$$G''(1) \geq c(c - 1).$$
 - c) Démontrer que cette inégalité ne peut devenir une égalité que lorsque c est un nombre entier. Quelle est la valeur minimale de $G''(1)$ dans le cas contraire?

Partie 4 : approximations polynomiales

Dans cette partie, on note $\| \cdot \|$ la norme de la convergence uniforme sur l'espace des fonctions réelles bornées sur le segment $[-1, +1]$.

1. Soit F l'espace vectoriel normé obtenu en munissant de la norme $\| \cdot \|$ l'espace des fonctions de classe C^∞ sur $[-1, +1]$.
 - a) L'ensemble des fonctions f de F qui vérifient $f(0) = 0$ est-il une partie fermée de F ?
 - b) Montrer que l'ensemble des fonctions f de F pour lesquelles $f(x)$ est négligeable devant x lorsque x tend vers 0 n'est pas une partie fermée de F .

Pour le prouver, on pourra utiliser la suite des fonctions $f_n : x \mapsto \sqrt{x^2 + \frac{1}{n}} - \sqrt{\frac{1}{n}}$.

2. On note T l'application qui, à chaque fonction f de F , associe la fonction $T(f)$ définie sur $[-1, +1]$ par :

$$\forall x \in [-1, +1], T(f)(x) = \int_0^x f(t) dt.$$

- a) Démontrer que T est une application continue de F dans F .
- b) Soit $f \in F$ et $n \in \mathbb{N}^*$. On note T^n la composée n -ième $T \circ T \circ \dots \circ T$ (n fois T). Démontrer que $T^n(f)$ est l'unique fonction de F nulle en 0 dont la dérivée n -ième est f et dont toutes les dérivées d'ordre inférieur à n s'annulent en 0.
- c) L'application T est-elle injective? Est-elle surjective?
3. Soit $k \in \mathbb{N}^*$ et f une fonction de F plate d'ordre k en 0.
- a) Justifier l'existence d'une suite $(P_n)_{n \in \mathbb{N}^*}$ de fonctions polynomiales convergeant uniformément vers $f^{(k+3)}$ sur le segment $[-1, +1]$.
- b) Démontrer que la suite $(T^{k+3}(P_n))_{n \in \mathbb{N}^*}$ converge uniformément sur $[-1, +1]$ vers une fonction de la forme $f + R$, où R est une fonction polynomiale de degré inférieur ou égal à $k + 2$.
- c) Justifier l'existence d'une suite de polynômes $(Q_n)_{n \in \mathbb{N}^*}$ vérifiant les propriétés suivantes :
- pour tout $n \in \mathbb{N}^*$, $Q_n(-1) = f(-1)$, $Q_n(0) = f(0)$ et $Q_n(1) = f(1)$;
 - pour tout $n \in \mathbb{N}^*$, Q_n est une fonction plate à l'ordre k en 0;
 - la suite $(Q_n)_{n \in \mathbb{N}^*}$ converge uniformément vers f sur $[-1, +1]$.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2016

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet se compose d'un problème et d'un exercice pouvant être traités de manière indépendante. Les différentes parties du problème peuvent également, dans une certaine mesure, être traitées de manière indépendante, quitte à admettre les résultats des parties précédentes. Une liste de résultats susceptibles d'être utiles, et pouvant être utilisés sans qu'il soit nécessaire de les redémontrer, est donnée au début.

La qualité de la rédaction et la précision des réponses apportées entrent pour une part importante dans la notation. Les réponses absurdes sont pénalisées.

Rappel de résultats

Ci-après figurent plusieurs résultats que l'on pourra utiliser sans qu'il soit nécessaire de les redémontrer. Par souci de simplicité, certains énoncés sont donnés dans des cas particuliers, et non pas dans le cadre le plus général possible.

1. **Inégalité triangulaire pour l'espérance** : pour toute variable aléatoire réelle bornée X , on a l'inégalité :

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$$

2. **Inégalité pour l'espérance d'une puissance** : étant donné une variable aléatoire réelle bornée X et un nombre réel $h \geq 1$, on a $(\mathbb{E}(|X|))^h \leq \mathbb{E}(|X|^h)$.
3. **Espérance conditionnelle à un événement** : étant donné une variable aléatoire réelle bornée X , et un événement A tel que $\mathbb{P}(A) > 0$, on a

$$\mathbb{E}(X|A) = \frac{\mathbb{E}(X \cdot \mathbf{1}_A)}{\mathbb{P}(A)},$$

où $\mathbf{1}_A$ est la variable aléatoire indicatrice de l'événement A , c'est-à-dire la variable aléatoire définie par :

$$\mathbf{1}_A = \begin{cases} 1 & \text{si } A \text{ est réalisé} \\ 0 & \text{si } A \text{ n'est pas réalisé} \end{cases}.$$

4. **Décomposition de l'espérance** : étant donné une variable aléatoire réelle bornée X , et un système complet d'événements $(A_i)_{i \in I}$ (où I est un ensemble fini ou dénombrable), on a :

$$\mathbb{E}(X) = \sum_{i \in I} \mathbb{E}(X|A_i) \cdot \mathbb{P}(A_i). \quad (1)$$

Problème

Un résultat fondamental en théorie des probabilités est le théorème central limite, dont l'une des formulations possibles est la suivante : étant donné une suite de variables aléatoires réelles $(X_n)_{n \geq 1}$ indépendantes et de même loi, satisfaisant en outre les conditions $\mathbb{E}(X_1) = 0$ et $\mathbb{V}(X_1) = 1$ (où la notation $\mathbb{V}$ désigne la variance), on a la convergence en loi

$$\frac{S_n}{\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y, \quad (2)$$

où

$$S_n = X_1 + \cdots + X_n,$$

et où Y suit la loi normale $\mathcal{N}(0, 1)$. On rappelle que (2) équivaut au fait que, pour tout $x \in \mathbb{R}$, on a

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du. \quad (3)$$

Dans ce problème, on se propose notamment d'établir le résultat plus élémentaire suivant : étant donné une suite de variables aléatoires $(X_n)_{n \geq 1}$ indépendantes et de même loi, satisfaisant en outre les trois conditions

- (a) $\mathbb{P}(|X_1| \leq c) = 1$ pour une certaine constante $c > 0$,
- (b) $\mathbb{E}(X_1) = 0$,
- (c) $\mathbb{V}(X_1) = 1$,

on a, pour tout entier $m \geq 1$, le fait que

$$\lim_{n \rightarrow +\infty} \mathbb{E}\left(\left(\frac{S_n}{\sqrt{n}}\right)^m\right) = \mathbb{E}(Y^m),$$

où, comme précédemment, $S_n = X_1 + \cdots + X_n$, et où Y suit la loi normale $\mathcal{N}(0, 1)$.

Dans la suite, on utilisera les notations

$$T_n = \frac{S_n}{\sqrt{n}}$$

et

$$M(m) = \mathbb{E}(Y^m).$$

Dans le cadre de ce problème, étant donné une fonction continue $f : \mathbb{R} \rightarrow \mathbb{R}$, on dira que l'intégrale $\int_{-\infty}^{+\infty} f(x) dx$ existe lorsqu'il existe un nombre réel I tel que

$$\lim_{A \rightarrow +\infty} \int_{-A}^{+A} f(x) dx = I,$$

et l'on écrira alors $\int_{-\infty}^{+\infty} f(x) dx = I$. (Il s'agit d'une définition simplifiée uniquement destinée à ce problème.)

Partie I – Calcul de $M(m)$

Dans cette partie, on calcule explicitement pour tout entier $m \geq 1$ la valeur de $M(m)$.

Question 1) Que représente la fonction $x \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ pour une variable aléatoire Y de loi normale $\mathcal{N}(0, 1)$? En déduire l'identité :

$$M(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} dx = 1. \quad (4)$$

Question 2) Montrer que, pour tout $m \geq 1$, on a l'identité

$$M(m) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^m e^{-x^2/2} dx.$$

Question 3) Montrer que, pour tout entier $k \geq 0$, on a $M(2k+1) = 0$.

Question 4) Montrer que, pour tout entier $k \geq 0$, on a $M(2k+2) = (2k+1) \cdot M(2k)$.

Indication : On pourra utiliser la formule d'intégration par parties.

Question 5) Dédire de ce qui précède l'identité, valable pour tout $k \geq 0$:

$$M(2k) = \frac{(2k)!}{2^k \cdot k!}.$$

Partie II – Un résultat sur les suites

Cette partie est consacrée à la preuve d'un résultat général sur les suites, qui est utilisé dans la partie III.

Etant donné une suite $(u_n)_{n \geq 1}$ de nombres réels, un nombre réel $q \geq 0$ et un nombre réel α tels que la propriété suivante soit vérifiée :

$$\lim_{n \rightarrow +\infty} \frac{u_{n+1} - u_n}{n^q} = \alpha, \quad (5)$$

on va prouver le fait que

$$\lim_{n \rightarrow +\infty} \frac{u_n}{n^{q+1}} = \frac{\alpha}{q+1}. \quad (6)$$

Question 6) Montrer que, pour tout nombre réel $q \geq 0$, on a

$$\lim_{n \rightarrow +\infty} \frac{\sum_{\ell=1}^{n-1} \ell^q}{n^{q+1}} = \frac{1}{q+1}.$$

Indication : On pourra d'abord établir que pour tout $1 \leq \ell \leq n-1$, on a un encadrement donné par $\int_{\ell-1}^{\ell} x^q dx \leq \ell^q \leq \int_{\ell}^{\ell+1} x^q dx$, et en déduire un encadrement de $\sum_{\ell=1}^{n-1} \ell^q$.

Question 7) On considère une suite $(v_n)_{n \geq 1}$ de nombres réels telle que $\lim_{n \rightarrow +\infty} v_n = 0$. Montrer que pour tout entier $q \geq 0$,

$$\lim_{n \rightarrow +\infty} \frac{\sum_{\ell=1}^{n-1} v_{\ell} \cdot \ell^q}{n^{q+1}} = 0.$$

Indication : si n_0 et $\epsilon > 0$ sont tels que, pour tout $n \geq n_0$, on a $|v_n| \leq \epsilon$, utiliser une majoration très simple de $\left| \sum_{\ell=n_0}^{n-1} v_{\ell} \cdot \ell^q \right|$ faisant intervenir $\sum_{\ell=n_0}^{n-1} \ell^q$.

Question 8) En faisant la synthèse des questions précédentes, établir le résultat sur les suites souhaité : la propriété (5) entraîne la propriété (6).

Partie III – Convergence de $\mathbb{E}(T_n^m)$ vers $M(m)$

Dans cette partie, on prouve le fait que, pour tout $m \geq 1$, on a

$$\lim_{n \rightarrow +\infty} \mathbb{E}(T_n^m) = M(m). \quad (7)$$

On commence par étudier les cas particuliers $m = 1, 2, 3$, avant d'effectuer une preuve dans le cas général.

Question 9) Montrer que, pour tout $n \geq 1$, on a $\mathbb{E}(S_n) = 0$, et en déduire (7) dans le cas particulier où $m = 1$.

Question 10) Montrer que, pour tout $n \geq 1$, on a $\mathbb{E}(S_n^2) = n$, et en déduire (7) dans le cas particulier où $m = 2$.

Indication : On pourra écrire que $S_{n+1}^2 = (S_n + X_{n+1})^2$, développer cette expression, et en déduire la valeur de $\mathbb{E}(S_{n+1}^2) - \mathbb{E}(S_n^2)$.

Question 11) Montrer que, pour tout $n \geq 1$, on a $\mathbb{E}(S_n^3) = n\mathbb{E}(X_1^3)$, et en déduire (7) dans le cas particulier où $m = 3$.

Indication : On pourra écrire que $S_{n+1}^3 = (S_n + X_{n+1})^3$, développer cette expression, et en déduire la valeur de $\mathbb{E}(S_{n+1}^3) - \mathbb{E}(S_n^3)$.

Question 12) On va maintenant prouver par récurrence le fait que, pour tout entier $q \geq 1$, on a la propriété (H_q) définie par

$$(H_q) : \lim_{n \rightarrow +\infty} \mathbb{E}(T_n^{2q}) = M(2q).$$

On suit pour cela les étapes indiquées ci-après.

- (i) Vérifier la propriété (H_q) pour $q = 1$.
- (ii) Montrer que, si l'on suppose que la propriété (H_q) est vérifiée pour un certain $q \geq 1$, on a, pour tout entier $k \leq 2q - 1$, le fait que

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}(|S_n|^k)}{n^q} = 0.$$

Indication : On pourra utiliser les points 1 et 2 du rappel de résultats.

- (iii) Montrer que, si l'on suppose que la propriété (H_q) est vérifiée pour un certain $q \geq 1$, on a

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{E}(S_{n+1}^{2(q+1)}) - \mathbb{E}(S_n^{2(q+1)})}{n^q} = (2q+1) \cdot (q+1) \cdot M(2q).$$

Indication : On pourra écrire que $S_{n+1}^{2(q+1)} = (S_n + X_{n+1})^{2(q+1)}$, développer cette expression, et penser à utiliser (ii) pour traiter le comportement d'une partie des termes.

- (iv) Montrer que, si l'on suppose que la propriété (H_q) est vérifiée pour un certain $q \geq 1$, la propriété (H_{q+1}) est également vérifiée.

Indication : Utiliser ce qui précède et les résultats des parties I et II.

Question 13) Montrer que, pour tout entier $q \geq 0$, on a la propriété

$$\lim_{n \rightarrow +\infty} \mathbb{E}(T_n^{2q+1}) = 0.$$

Indication : Adapter les arguments précédents.

Question 14) Effectuer un bilan des résultats précédents, établissant la validité de la propriété (7) pour tout $m \geq 1$.

Partie IV – Convergence de $\mathbb{E}(f(T_n))$ pour f continue et bornée

Une formulation mathématiquement équivalente de l'énoncé (3) du théorème central limite est que, pour toute fonction **continue et bornée** $f : \mathbb{R} \rightarrow \mathbb{R}$, on a la convergence :

$$\lim_{n \rightarrow +\infty} \mathbb{E}(f(T_n)) = \mathbb{E}(f(Y)), \quad (8)$$

où Y suit la loi normale $\mathcal{N}(0, 1)$. Cette partie est consacrée à établir le fait que la propriété (8) entraîne effectivement la propriété (3).

Question 15) Montrer que, pour tout $x \in \mathbb{R}$, on a $\mathbb{P}(T_n \leq x) = \mathbb{E}(g_x(T_n))$, où g_x est définie par

$$g_x(t) = \begin{cases} 1 & \text{si } t \leq x \\ 0 & \text{si } t > x \end{cases}.$$

La fonction g_x ainsi définie est-elle bornée? Est-elle continue?

Question 16) Etant donné $\epsilon > 0$, définir une fonction **continue** $f_{x,\epsilon} : \mathbb{R} \rightarrow \mathbb{R}$ telle que l'on ait :

- $g_x(t) \leq f_{x,\epsilon}(t)$ pour tout $t \in \mathbb{R}$,
- $f_{x,\epsilon}(t) \in [0, 1]$ pour tout $t \in \mathbb{R}$,
- $g_x(t) = f_{x,\epsilon}(t)$ pour tout $t \notin]x, x + \epsilon[$.

Question 17) Montrer que, étant donné $x \in \mathbb{R}$, et une variable aléatoire Y suivant la loi normale $\mathcal{N}(0, 1)$, on doit avoir, pour la fonction $f_{x,\epsilon}$ définie à la question 16), la propriété :

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}(f_{x,\epsilon}(Y)) = \mathbb{E}(g_x(Y)).$$

Question 18) Dédurre de ce qui précède que, si l'on suppose la validité de (8) pour toute fonction continue bornée f , alors, pour tout $\alpha > 0$, on peut trouver un entier n_0 tel que, pour tout $n \geq n_0$, on a :

$$\mathbb{E}(g_x(T_n)) \leq \mathbb{E}(g_x(Y)) + \alpha.$$

Question 19) Adapter le raisonnement des questions 16) à 18) pour établir que, si l'on suppose la validité de (8) pour toute fonction continue bornée f , alors, pour tout $\alpha > 0$, on peut trouver un entier n_0 tel que, pour tout $n \geq n_0$, on a :

$$\mathbb{E}(g_x(T_n)) \geq \mathbb{E}(g_x(Y)) - \alpha,$$

où Y suit la loi normale $\mathcal{N}(0, 1)$.

Question 20) Conclure de ce qui précède le fait que, si l'on suppose la validité de (8) pour toute fonction continue bornée f , on a la propriété (3).

Partie V – Un petit exemple d'application du théorème central limite

Sur une période d'une année, un assureur prend à sa charge les montants des sinistres affectant une population de 10000 assurés. En modélisant les montants individuels des sinistres associés à chacun des assurés par une famille de variables aléatoires $Z_1, \dots, Z_{10000}$ indépendantes et de même loi, la somme $Z_1 + \dots + Z_{10000}$ représente donc le montant total à la charge de l'assureur sur une période d'une année.

Question 21) En supposant que les variables Z_i (exprimées en euros) vérifient $\mathbb{E}(Z_i) = 4$ et $\mathbb{V}(Z_i) = 4$, utiliser le théorème central limite (voir (3)) pour proposer une approximation des probabilités suivantes :

$$p_1 = \mathbb{P}(Z_1 + \dots + Z_{10000} \geq 40200) \text{ et } p_2 = \mathbb{P}(Z_1 + \dots + Z_{10000} \geq 40400).$$

Données : valeurs approchées à $\pm 10^{-3}$ près de $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$ pour certaines valeurs de x .

x	-2	-1,5	-1	-0,5	0	0,5	1	1,5	2
$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$	0,023	0,067	0,159	0,309	0,5	0,691	0,841	0,933	0,977

Indication : On pourra transformer les variables Z_i en des variables X_i de manière à garantir les conditions $\mathbb{E}(X_i) = 0$ et $\mathbb{V}(X_i) = 1$.

Exercice

On considère une population d'individus évoluant au cours du temps. L'écoulement du temps est représenté par une succession de dates $t = \dots, -2, -1, 0, 1, 2, \dots$, la date $t = 0$ étant prise comme origine pour l'étude de la population. Un individu présent dans la population à la date t et né à la date $s \leq t$ est considéré comme ayant un âge égal à $t - s$ à la date t . Pour tout entier $t \geq 0$, et tout entier $x \geq 0$, on note $N_x(t)$ le nombre d'individus présents dans la population à la date t et ayant un âge égal à x . On suppose l'existence d'un âge maximal atteignable $x^* \geq 1$, si bien que $N_x(t)$ est automatiquement égal à 0 lorsque $x > x^*$. La **durée de vie** d'un individu est définie comme le nombre total de périodes de temps pendant lesquelles l'individu s'est trouvé en vie. Ainsi, un individu né à la date t_1 , encore en vie à la date $t_2 \geq t_1$, et décédé à la date $t_2 + 1$, a une durée de vie égale à $t_2 - t_1$. Finalement, on note $N(t)$ le nombre **total** (tous âges confondus) d'individus présents dans la population à la date t .

Le processus d'évolution de la population au cours du temps est modélisé de la manière suivante : les nombres $N_0(x)$ pour $x = 0, \dots, x^*$ (composition initiale de la population) sont supposés connus. Ensuite, pour tout $t \geq 0$, étant donné l'évolution de la population jusqu'à la date t , et indépendamment de tous les autres individus de la population, la probabilité pour un individu d'âge x à la date t d'être décédé à la date $t + 1$ est donnée par un nombre $q_x > 0$. On a $q_{x^*} = 1$. Un individu décédé à la date $t + 1$ n'est alors plus considéré comme faisant partie de la population. D'autre part, on suppose que les nombres de nouveaux-nés $N_0(1), N_0(2), \dots$ forment une suite de variables aléatoires indépendantes entre elles, et indépendantes de la survie/non-survie des autres individus présents dans la population.

Question A) Pour $t \geq 0$, exprimer $N(t)$ en fonction de $N_0(t), \dots, N_{x^*}(t)$.

Question B) Etant donné $x \geq 0$, un entier $n \geq 0$, et un entier $t \geq 0$, exprimer la valeur de $\mathbb{E}(N_{x+1}(t+1) | N_x(t) = n)$ en fonction de n et de q_x .

Question C) En déduire une expression de $\mathbb{E}(N_{x+1}(t+1))$ en fonction de $\mathbb{E}(N_x(t))$ et de q_x .

Question D) On considère un individu d'âge $y \geq 0$ à la date 0. On note T la variable aléatoire correspondant à la durée de vie (totale) de cet individu, et S la variable aléatoire correspondant à la durée de vie **restante** de cet individu après la date 0, c'est-à-dire $S = T - y$. Donner une expression explicite de $\mathbb{P}(S = z)$ pour tout $z \geq 0$ en fonction des q_x .

Question E) Déduire du résultat de la question précédente une expression pour $\mathbb{E}(S)$ et pour $\mathbb{V}(S)$.

Question F) On dit que la population est **stationnaire en moyenne** lorsque, pour tout $t \geq 0$, et tout $x = 0, \dots, x^*$, on a $\mathbb{E}(N_x(t+1)) = \mathbb{E}(N_x(t))$.

Montrer qu'une condition nécessaire et suffisante pour que la population soit stationnaire en moyenne est que les équations suivantes soient satisfaites :

(i) pour tout $t \geq 1$, $\mathbb{E}(N_0(t)) = N_0(0)$,

(ii) pour tout $x = 0, \dots, x^* - 1$, on a $(1 - q_x)N_x(0) = N_{x+1}(0)$.

Question G) On suppose que la condition (i) de la question précédente est vérifiée. Montrer que, pour une valeur donnée de $N(0)$ et des q_x , il existe un et un seul jeu de valeurs $N_0(0), \dots, N_0(x^*)$ tel que la population soit stationnaire en moyenne. Donner une expression explicite de ces valeurs de $N_0(0), \dots, N_0(x^*)$, en fonction de $N(0)$ et des q_x .

Question H) On suppose que la population est stationnaire en moyenne. On numérote les individus présents dans la population à la date 0 par $i = 1, \dots, N(0)$, et l'on utilise la notation S_i pour désigner la durée de vie restante (après la date 0) de l'individu numéro i . Montrer que l'on a, pour tout $x \geq 0$, la relation : $N_x(0) = \sum_{i=1}^{N(0)} \mathbb{P}(S_i = x)$.

Question I) Justifier l'interprétation suivante de la relation obtenue en H) : « dans une population stationnaire en moyenne, le nombre moyen d'individus âgés de x années est égal au nombre moyen d'individus auxquels il reste exactement x années à vivre ».

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2017

Épreuve de mathématiques

Durée : 4h

Toutes les variables aléatoires qui interviennent dans ce problème sont définies sur le même espace probabilisable $(\Omega, \mathcal{A})$. On note $\mathbf{P}$ une probabilité sur cet espace. On note $\mathbf{E}(X)$ et $\mathbf{Var}(X)$ les espérance et variance (pour la probabilité $\mathbf{P}$) d'une variable X . Si $A \in \mathcal{A}$, on note $\mathbf{1}_A$ son indicatrice, c'est-à-dire l'application qui à chaque $\omega \in \Omega$ associe 1 si $\omega \in A$ et 0 sinon. Si $A \in \mathcal{A}$ est de probabilité non nulle, on note $\mathbf{P}_A$ la probabilité conditionnelle sachant A et, s'il y a lieu, $\mathbf{E}_A(X)$ l'espérance, pour $\mathbf{P}_A$, d'une variable aléatoire X .

On pose, pour tout entier naturel n non nul, $H_n = \sum_{k=1}^n \frac{1}{k}$.

L'objet du problème est, principalement, l'étude de la variable aléatoire égale au maximum de n variables indépendantes toutes de même loi géométrique.

La partie **I** regroupe des questions indépendantes dont les résultats seront utilisés par la suite.

Partie I. Préliminaires

1. Un résultat bien connu

- a) Montrer que la suite de terme général $u_n = H_n - \ln n$ est monotone.
 b) En déduire l'existence d'un réel noté γ pour lequel on a, quand l'entier n tend vers $+\infty$,

$$H_n = \ln n + \gamma + o(1).$$

2. Un résultat de bornitude...

Soit f une fonction continue de $\mathbb{R}^+$ dans $\mathbb{R}$. On suppose que la fonction f a une limite nulle en $+\infty$. Montrer que la fonction f est bornée.

3. Une formulation intégrale des moments d'une variable positive

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

- a) Établir, pour tout entier naturel N , l'égalité :

$$\sum_{k=0}^N \mathbf{P}([X > k]) = \sum_{k=1}^N k \mathbf{P}([X = k]) + (N+1) \mathbf{P}([X > N]).$$

- b) On suppose que la variable X possède une espérance.

(i) Établir l'égalité : $\lim_{N \rightarrow +\infty} (N+1) \mathbf{P}([X > N]) = 0$.

(ii) En déduire que la série $\sum \mathbf{P}([X > n])$ converge et qu'on a l'égalité :

$$\mathbf{E}(X) = \sum_{n=0}^{+\infty} \mathbf{P}([X > n]).$$

- c) On suppose, réciproquement, que la série $\sum \mathbf{P}([X > n])$ converge. Montrer que la variable X possède une espérance qu'on a l'égalité : $\mathbf{E}(X) = \sum_{n=0}^{+\infty} \mathbf{P}([X > n])$.

- d) (i) Vérifier que la fonction qui à chaque réel t associe $\mathbf{P}([X > t])$ est continue par morceaux sur $\mathbb{R}$.

(ii) Montrer que X a une espérance si, et seulement si, l'intégrale

$$\int_0^{+\infty} \mathbf{P}([X > t]) dt$$

converge et que, dans ce cas, on a l'égalité : $\mathbf{E}(X) = \int_0^{+\infty} \mathbf{P}([X > t]) dt$.

- e) Montrer que X a un moment d'ordre 2 si, et seulement si, l'intégrale

$$\int_0^{+\infty} t \mathbf{P}([X > t]) dt$$

converge et que, dans ce cas, on a l'égalité : $\mathbf{E}(X^2) = 2 \int_0^{+\infty} t \mathbf{P}([X > t]) dt$.

4. Autour du produit de convolution

Soit $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles. On pose, pour tout entier naturel n ,

$$w_n = \sum_{k=0}^n u_k v_{n-k}.$$

On suppose que, pour tout réel x , les séries $\sum u_n x^n$ et $\sum v_n x^n$ convergent absolument. On rappelle, ou on admet, qu'alors, pour tout réel x , la série $\sum w_n x^n$ converge absolument et qu'on a l'égalité :

$$\sum_{n=0}^{+\infty} w_n x^n = \left(\sum_{n=0}^{+\infty} u_n x^n \right) \left(\sum_{n=0}^{+\infty} v_n x^n \right).$$

Soit $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ deux suites réelles. On pose, pour tout entier naturel n ,

$$c_n = \sum_{k=0}^n \binom{n}{k} a_k b_{n-k}.$$

On suppose que, pour tout réel x , les séries $\sum \frac{a_n}{n!} x^n$ et $\sum \frac{b_n}{n!} x^n$ convergent absolument. Montrer que, pour tout réel x , la série $\sum \frac{c_n}{n!} x^n$ converge absolument et qu'on a l'égalité :

$$\sum_{n=0}^{+\infty} \frac{c_n}{n!} x^n = \left(\sum_{n=0}^{+\infty} \frac{a_n}{n!} x^n \right) \left(\sum_{n=0}^{+\infty} \frac{b_n}{n!} x^n \right).$$

5. Loi conditionnelle d'un vecteur aléatoire

- a) Soit X une variable aléatoire suivant la loi géométrique de paramètre $p \in]0, 1[$. Quelle est la loi de $X - 1$ pour la probabilité conditionnelle $\mathbf{P}_{[X > 1]}$?
- b) Soit $r \in \mathbb{N}^*$ et $X_1, X_2, \dots, X_r$ des variables aléatoires **indépendantes** toutes de même loi, chacune suivant la loi géométrique (sur $\mathbb{N}^*$) de paramètre $p \in]0, 1[$. On note

$$A = \bigcap_{i=1}^r [X_i > 1].$$

Montrer que le vecteur aléatoire $(X_1 - 1, X_2 - 1, \dots, X_r - 1)$ a, pour la probabilité conditionnelle $\mathbf{P}_A$, même loi que le vecteur aléatoire $(X_1, X_2, \dots, X_r)$ (pour la probabilité $\mathbf{P}$).

6. La formule de l'espérance totale

Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé.

- a) Soit A un événement de probabilité non nulle et soit X une variable aléatoire possédant une espérance. Montrer que le produit $X \mathbf{1}_A$ a une espérance et exprimer $\mathbf{E}(X \mathbf{1}_A)$ à l'aide de $\mathbf{E}_A(X)$.
- b) On considère des événements $A_1, A_2, \dots, A_n$, chacun étant de probabilité non nulle, formant une partition de Ω . Soit X une variable aléatoire, définie sur l'espace $(\Omega, \mathcal{A}, \mathbf{P})$, à valeurs dans $\mathbb{N}$ et possédant une espérance. Établir l'égalité :

$$\mathbf{E}(X) = \sum_{k=1}^n \mathbf{E}_{A_k}(X) \mathbf{P}(A_k).$$

Dans toute la suite du problème on considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires **in-dépendantes** toutes de même loi, chacune suivant la loi géométrique (sur $\mathbb{N}^*$) de paramètre $p \in]0, 1[$. On pose $q = 1 - p$.

On note, pour tout entier naturel n non nul et tout $\omega \in \Omega$,

$$I_n(\omega) = \min(X_1(\omega), X_2(\omega), \dots, X_n(\omega)) \quad \text{et} \quad M_n(\omega) = \max(X_1(\omega), X_2(\omega), \dots, X_n(\omega)).$$

7. Montrer que, pour tout entier naturel n non nul, l'application I_n est une variable aléatoire sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$. Il en est de même, et on l'admet, de l'application M_n .

Partie II. Étude du minimum

1. Montrer que, pour tout entier naturel n non nul, I_n suit une loi géométrique dont on donnera le paramètre.
2. Quelles sont les limites de $\mathbf{E}(I_n)$, $\mathbf{Var}(I_n)$ et, pour tout entier naturel k non nul, $\mathbf{P}([I_n = k])$ lorsque l'entier n tend vers $+\infty$?
3. a) Soit $\omega \in \Omega$. Justifier l'existence d'une limite finie, notée $\ell(\omega)$, pour la suite de terme général $I_n(\omega)$.
b) Exprimer la partie $\mathcal{L}$ définie par $\mathcal{L} = \{\omega \in \Omega; \ell(\omega) = 1\}$ en fonction des événements $[I_n = 1]$ et en déduire que la partie $\mathcal{L}$ est un événement presque sûr.

Partie III. Étude du maximum

1. L'espérance de M_n tend vers $+\infty$
a) Soit $n \in \mathbb{N}^*$.
(i) Justifier l'existence d'une espérance pour la variable M_n et l'encadrement

$$\frac{1}{p} \leq \mathbf{E}(M_n) \leq \frac{n}{p}.$$

(ii) Déterminer, pour tout entier naturel k , la valeur de $\mathbf{P}([M_n \leq k])$.

- b) (i) Soit $K \in \mathbb{N}^*$. Établir, pour tout entier naturel n non nul, l'inégalité :

$$\mathbf{E}(M_n) \geq \mathbf{E}(M_n \mathbf{1}_{[M_n \leq K]}) + K \mathbf{P}([M_n > K]).$$

(ii) En déduire que $\lim_{n \rightarrow +\infty} \mathbf{E}(M_n) = +\infty$.

2. Deux expressions de l'espérance de M_n

Soit $n \in \mathbb{N}^*$.

a) Établir l'égalité : $\mathbf{E}(M_n) = \sum_{i=1}^n \binom{n}{i} (-1)^{i-1} \frac{1}{1-q^i}$. On utilisera le résultat de I-3-b-ii).

b) On note $[x]$ la partie entière d'un réel x . Établir l'égalité :

$$\mathbf{E}(M_n) = \int_0^{+\infty} \left(1 - (1 - q^{[t]})^n\right) dt.$$

On utilisera le résultat de I-3-d-ii).

3. Estimation de l'espérance de M_n

a) Justifier, pour tout entier naturel k , la convergence de l'intégrale $\int_0^{+\infty} q^t (1 - q^t)^k dt$ et déterminer sa valeur.

b) En déduire, pour tout entier naturel n non nul, l'égalité :

$$\int_0^{+\infty} \left(1 - (1 - q^t)^n\right) dt = -\frac{H_n}{\ln q}.$$

c) (i) Établir, pour tout entier naturel n non nul, l'encadrement :

$$-\frac{H_n}{\ln q} \leq \mathbf{E}(M_n) \leq -\frac{H_n}{\ln q} + 1.$$

En déduire l'équivalent de $\mathbf{E}(M_n) \sim -\frac{\ln n}{\ln q}$ quand l'entier n tend vers $+\infty$.

(ii) **Déduire** de l'encadrement précédent, pour tout $x \in]0, 1[$, l'encadrement :

$$-\frac{x}{1-x} \leq \ln(1-x) \leq -x.$$

4. Estimation de la variance de M_n

a) Soit $k \in \mathbb{N}$.

(i) Justifier la convergence de l'intégrale $\int_0^{+\infty} t q^t (1 - q^t)^k dt$ dont on note α_k la valeur.

(ii) Établir l'égalité $\alpha_k = -\frac{1}{(k+1)\ln q} \int_0^{+\infty} \left(1 - (1 - q^t)^{k+1}\right) dt$.

b) Établir, pour tout entier naturel n non nul, l'égalité $\int_0^{+\infty} t \left(1 - (1 - q^t)^n\right) dt = \frac{1}{\ln^2 q} \sum_{k=1}^n \frac{H_k}{k}$ puis l'encadrement

$$\frac{2}{\ln^2 q} \sum_{k=1}^n \frac{H_k}{k} \leq \mathbf{E}(M_n^2) \leq 1 + \frac{2}{\ln^2 q} \sum_{k=1}^n \frac{H_k}{k} - 2 \frac{H_n}{\ln q}.$$

c) Établir, pour tout entier naturel n non nul, l'égalité : $H_n^2 = 2 \sum_{k=1}^n \frac{H_k}{k} - \sum_{k=1}^n \frac{1}{k^2}$.

d) En déduire l'estimation asymptotique $\mathbf{Var}(M_n) = O(\ln n)$ quand l'entier n tend vers $+\infty$.

Partie IV. Comportement asymptotique de la suite $(M_n)_{n \in \mathbb{N}^*}$

1. Prouver, pour tout $x \in [0, 1[$, l'encadrement : $-\frac{x^2}{1-x} \leq x + \ln(1-x) \leq 0$.

2. Soit $n \in \mathbb{N}^*$. Prouver, pour tout entier naturel k non nul, les majorations :

$$\left| e^{-nq^k} - (1 - q^k)^n \right| \leq \frac{n q^{2k}}{1 - q^k} e^{-nq^k} \leq \frac{n q^{2k}}{1 - q} e^{-nq^k}.$$

3. Montrer que la fonction $x \mapsto x^2 e^{-x}$ est bornée sur $\mathbb{R}^+$.

4. Déduire des questions précédentes le résultat asymptotique suivant :

$$\lim_{n \rightarrow \infty} \sup_{k \in \mathbb{N}^*} \left| \mathbf{P}([M_n \leq k]) - e^{-nq^k} \right| = 0.$$

Partie V. Où l'on retrouve une formule exacte pour l'espérance de M_n

Dans cette dernière partie on suppose que $p = \frac{1}{2}$ et on considère un entier $n \in \mathbb{N}^*$.

On pose $m_0 = 0$ et, pour tout entier naturel n non nul, $m_n = \mathbf{E}(M_n)$.

1. On note $B_0 = \bigcap_{i=1}^n [X_i > 1]$, $B_n = \bigcap_{i=1}^n [X_i = 1]$ et, pour tout $k \in \llbracket 1, n-1 \rrbracket$,

$$B_k = \left(\bigcap_{i=1}^k [X_i = 1] \right) \cap \left(\bigcap_{i=k+1}^n [X_i > 1] \right).$$

- a) Établir, pour tout $k \in \llbracket 0, n \rrbracket$, l'égalité : $\mathbf{E}_{B_k}(M_n) = 1 + m_{n-k}$.
 b) En utilisant la formule de l'espérance totale, établir, pour tout entier naturel n non nul, l'égalité :

$$m_n = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} (1 + m_{n-k}).$$

2. Justifier, pour tout réel x , la convergence de la série de terme général $\frac{m_n}{n!} x^n$.

On note, pour tout réel x , $M(x) = \sum_{n=0}^{+\infty} \frac{m_n}{n!} x^n$.

3. Établir, pour tout réel x , l'égalité : $M(2x) = e^{2x} - 1 + M(x) e^x$.

On note, pour tout réel x , $G(x) = M(x) e^{-x}$.

On a donc, pour tout réel x , $G(2x) = 1 - e^{-2x} + G(x)$.

4. Établir, pour tout entier naturel n non nul, l'égalité :

$$\mathbf{E}(M_n) = m_n = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \frac{2^k}{2^k - 1}.$$

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2017

Épreuve à option (A) : Mathématiques

Durée : 4h

Dans tout le problème, n désigne un nombre entier supérieur ou égal à deux.

Selon une formule classique, que l'on pourra admettre dans toute la suite, toute matrice de $\mathcal{M}_n(\mathbb{R})$ de la forme

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & & x_n^{n-1} \end{pmatrix}$$

admet pour déterminant le nombre réel $\prod_{1 \leq i < j \leq n} (x_j - x_i)$, appelé *déterminant de Vandermonde*, que l'on notera dans la suite $v_n(x_1, x_2, \dots, x_n)$.

L'objet du problème est l'étude de déterminants de matrices d'une forme plus générale, appelés *déterminants vandermondiens*, définis comme suit.

Pour toute famille $(f_1, f_2, \dots, f_n)$ de fonctions définies sur un ensemble $\mathcal{S}$, à valeurs réelles, et tout élément $(x_1, x_2, \dots, x_n)$ de $\mathcal{S}^n$, on note

$$V_{f_1, \dots, f_n}(x_1, \dots, x_n) = \begin{pmatrix} f_1(x_1) & f_1(x_2) & \dots & f_1(x_n) \\ f_2(x_1) & f_2(x_2) & \dots & f_2(x_n) \\ \vdots & \vdots & & \vdots \\ f_n(x_1) & f_n(x_2) & \dots & f_n(x_n) \end{pmatrix}$$

et $v_{f_1, \dots, f_n}(x_1, \dots, x_n)$ le déterminant de cette matrice.

Bien que les déterminants vandermondiens interviennent dans les trois parties du sujet, celles-ci sont largement indépendantes.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Partie 1 : polynômes de Tchebycheff de seconde espèce

Dans cette partie, $x_1, x_2, \dots, x_n$ désignent n nombres réels vérifiant les inégalités strictes

$$0 < x_1 < x_2 < \dots < x_n < \pi$$

et on note $s_1, s_2, \dots, s_n$ les n fonctions définies par :

$$\forall i \in \llbracket 1, n \rrbracket, \forall x \in \mathbb{R}, \quad s_i(x) = \sin(ix).$$

On considère la suite $(U_p(X))_{p \in \mathbb{N}}$ des polynômes de $\mathbb{R}[X]$ définis par :

$$\begin{cases} U_0(X) = 1 \\ U_1(X) = 2X \\ \forall p \in \mathbb{N}, \quad U_{p+2}(X) = 2XU_{p+1}(X) - U_p(X) \end{cases}.$$

1.
 - a) Calculer $U_2(X)$ et $U_3(X)$, puis indiquer la parité de la fonction $x \mapsto U_p(x)$, selon les valeurs de p .
 - b) Pour tout entier naturel p , déterminer le degré et le coefficient dominant du polynôme $U_p(X)$.
 - c) Justifier que $(U_0(X), U_1(X), \dots, U_n(X))$ est une base de $\mathbb{R}_n[X]$.
2. Soit $t \in \mathbb{R}$.
 - a) Pour $p \in \mathbb{N}$, exprimer $\sin((p+2)t) + \sin(pt)$ en fonction de $\sin((p+1)t)$ et $\cos(t)$.
 - b) En déduire que pour tout entier naturel p , on a :

$$\sin((p+1)t) = \sin(t) U_p(\cos(t)).$$

- c) Justifier que le polynôme $U_n(X)$ possède n racines distinctes, qui appartiennent toutes à l'intervalle ouvert $] -1, +1 [$.
3. Pour tout couple $(i, j) \in \llbracket 1, n \rrbracket^2$, on pose : $c_{i,j} = 2^{1-i} U_{i-1}(\cos(x_j))$.
On note C la matrice $(c_{i,j})_{(i,j) \in \llbracket 1, n \rrbracket^2}$ et $L_1, L_2, \dots, L_n$ les lignes de cette matrice.
 - a) Justifier l'existence de nombres réels $\lambda_1, \dots, \lambda_{n-1}$ tels que :

$$L_n - \sum_{j=1}^{n-1} \lambda_j L_j = (\cos^{n-1}(x_1) \quad \cos^{n-1}(x_2) \quad \dots \quad \cos^{n-1}(x_n)).$$

- b) Démontrer que le déterminant de C est égal à $v_n(\cos(x_1), \cos(x_2), \dots, \cos(x_n))$.
 - c) Établir la formule :

$$v_{s_1, \dots, s_n}(x_1, \dots, x_n) = 2^{n(n-1)/2} \left(\prod_{k=1}^n \sin(x_k) \right) \left(\prod_{1 \leq i < j \leq n} (\cos(x_j) - \cos(x_i)) \right).$$

- d) En déduire l'indépendance linéaire des fonctions $s_1, s_2, \dots, s_n$.

4. a) Calculer, pour tout couple (p, q) de nombres entiers naturels, l'intégrale

$$\int_0^\pi \sin(pt) \sin(qt) dt.$$

- b) En déduire que :

$$\forall (p, q) \in \mathbb{N}^2, \quad \frac{2}{\pi} \int_{-1}^{+1} U_p(t) U_q(t) \sqrt{1-t^2} dt = \begin{cases} 1 & \text{si } p = q \\ 0 & \text{si } p \neq q \end{cases}.$$

5. On note V^* la comatrice de la matrice $V_{U_0, U_1, \dots, U_{n-1}}(\cos(x_1), \cos(x_2), \dots, \cos(x_n))$ et $v^*(i, j)$ ses coefficients (pour tout couple $(i, j) \in \llbracket 1, n \rrbracket^2$, $v^*(i, j)$ est le coefficient situé dans la i -ème ligne et la j -ème colonne de V^*).

- a) Justifier la continuité sur $\mathbb{R}$ de la fonction

$$\varphi : t \longmapsto v_{U_0, U_1, \dots, U_{n-1}}(\cos(x_1), \dots, \cos(x_{n-1}), t).$$

- b) Démontrer l'égalité

$$\frac{2}{\pi} \int_{-1}^{+1} (\varphi(t))^2 \sqrt{1-t^2} dt = \sum_{i=1}^n (v^*(i, n))^2.$$

Partie 2 : factorisations de déterminants vandermondiens

1. a) Justifier que la fonction $x \longmapsto \frac{\sin x}{x}$ admet un unique prolongement continu à $\mathbb{R}$ et démontrer, en utilisant une série entière, que ce prolongement est de classe C^∞ sur $\mathbb{R}$.

La fonction ainsi prolongée sera notée sinc (pour *sinus cardinal*).

b) Pour tout couple $(x, y) \in \mathbb{R}^2$, on note :
$$c(x, y) = \begin{cases} \frac{\cos(x) - \cos(y)}{x - y} & \text{si } x \neq y \\ -\sin(x) & \text{si } x = y \end{cases}.$$

Exprimer $c(x, y)$ à l'aide des fonctions sin et sinc, et en déduire que la fonction c est de classe C^∞ sur $\mathbb{R}^2$.

c) En déduire, à l'aide de la formule démontrée dans la question 3.c de la première partie, que la fonction $b : (x_1, x_2, \dots, x_n) \longmapsto \frac{v_{s_1, \dots, s_n}(x_1, \dots, x_n)}{v_n(x_1, \dots, x_n)}$ admet un prolongement de classe C^∞ à $\mathbb{R}^n$.

2. Dans cette question, f et g désignent des fonctions à valeurs réelles, continues sur un segment $[a, b]$ de $\mathbb{R}$ (avec $a < b$), dérivables sur l'intervalle ouvert $]a, b[$.

a) En appliquant le théorème de Rolle à une fonction de la forme $x \longmapsto v_{f_1, f_2, f_3}(a, b, x)$, justifier l'existence d'un élément c de $]a, b[$ vérifiant :

$$\frac{g(b)f(a) - g(a)f(b)}{b-a} = g'(c) \frac{bf(a) - af(b)}{b-a} - f'(c) \frac{bg(a) - ag(b)}{b-a}.$$

b) En quoi le résultat précédent est-il une généralisation du théorème des accroissements finis ?

3. Dans cette question f et g désignent des fonctions polynomiales à coefficients réels.

- La première est de degré n : $\forall x \in \mathbb{R}, f(x) = \sum_{i=0}^n a_i x^i$, avec $a_n \neq 0$.
- La seconde est de degré inférieur ou égal à n : $\forall x \in \mathbb{R}, b(x) = \sum_{i=0}^n b_i x^i$.

a) Soit k et ℓ deux éléments distincts de $\llbracket 1, n \rrbracket$.

Justifier que, pour $x \neq y$, le quotient $\frac{x^k y^\ell - x^\ell y^k}{x - y}$ est une combinaison linéaire de monômes de la forme $x^i y^j$, avec $(i, j) \in \llbracket 0, n-1 \rrbracket^2$.

b) Pour tout réel x , on note $C(x)$ la matrice-colonne $\begin{pmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^{n-1} \end{pmatrix}$ et $L(x)$ sa transposée.

Justifier l'existence d'une unique matrice carrée $B(f, g) \in \mathcal{M}_n(\mathbb{R})$ vérifiant, pour tout couple (x, y) de réels distincts, l'égalité :

$$\frac{f(x)g(y) - g(x)f(y)}{x - y} = L(x) B(f, g) C(y).$$

c) Démontrer que si une matrice-colonne M vérifie $L(x) M = 0$ pour tout réel x , alors M est nulle.

d) En déduire que si les deux polynômes $\sum_{i=0}^n a_i X^i$ et $\sum_{i=0}^n b_i X^i$ ont une racine réelle commune, alors la matrice $B(f, g)$ n'est pas inversible.

Partie 3 : unisolvance et interpolation

Soit $\mathcal{S}$ un ensemble non vide et F un sous-espace vectoriel de dimension n du $\mathbb{R}$ -espace vectoriel des applications de $\mathcal{S}$ dans $\mathbb{R}$.

On dit qu'un n -uplet $(x_1, x_2, \dots, x_n)$ d'éléments de $\mathcal{S}$ est *unisolvant* pour F si le seul élément de F qui s'annule en chacun des x_j est la fonction nulle :

$$\forall f \in F, \left[(\forall j \in \llbracket 1, n \rrbracket, f(x_j) = 0) \implies (\forall x \in \mathcal{S}, f(x) = 0) \right].$$

1. a) Justifier que $\mathcal{S}$ possède au moins n éléments.

b) Démontrer que, si un n -uplet $(x_1, x_2, \dots, x_n)$ d'éléments de $\mathcal{S}$ est unisolvant pour F , alors $x_1, x_2, \dots, x_n$ sont nécessairement deux à deux distincts.

2. Soit $(f_1, f_2, \dots, f_n)$ une base de F et $(x_1, x_2, \dots, x_n) \in \mathcal{S}^n$.

a) Démontrer que $(x_1, x_2, \dots, x_n)$ est unisolvant pour F si, et seulement si, le déterminant vandermondien $v_{f_1, \dots, f_n}(x_1, \dots, x_n)$ est différent de 0.

b) Soit $(y_1, y_2, \dots, y_n) \in \mathbb{R}^n$. Justifier que si $(x_1, x_2, \dots, x_n)$ est unisolvant pour F , alors il existe un unique élément f de F vérifiant :

$$\forall j \in \llbracket 1, n \rrbracket, f(x_j) = y_j.$$

c) En utilisant la formule démontrée dans la question 3.c de la première partie, justifier que pour tout n -uplet $(x_1, x_2, \dots, x_n)$ d'éléments deux à deux distincts de $]0, \pi[$ et toute fonction réelle g définie sur cet intervalle, il existe un unique vecteur $(a_1, a_2, \dots, a_n)$ vérifiant :

$$\forall j \in \llbracket 1, n \rrbracket, \sum_{i=1}^n a_i s_i(x_j) = g(x_j).$$

3. On dit que F est *unisolvant* si tout n -uplet $(x_1, x_2, \dots, x_n)$ d'éléments deux à deux distincts de $\mathcal{S}$ est unisolvant pour F .

a) Dans le cas où $\mathcal{S} = \mathbb{R}$, justifier que l'espace vectoriel des fonctions polynomiales de degré inférieur ou égal à $n - 1$ est unisolvant.

b) Dans le cas où $\mathcal{S} = \mathbb{R}$, montrer que l'espace vectoriel des fonctions polynomiales de degré inférieur ou égal à n qui s'annulent en 0 n'est pas unisolvant et déterminer les n -uplets $(x_1, x_2, \dots, x_n)$ qui sont unisolvants pour cet espace vectoriel.

4. Dans cette question, on étudie le cas où $\mathcal{S} = \mathbb{R}^2$ et on suppose que tous les éléments de l'espace vectoriel F sont des applications continues (de $\mathbb{R}^2$ dans $\mathbb{R}$).

On cherche à montrer que F ne peut pas être unisolvant.

Soit $(f_1, f_2, \dots, f_n)$ une base de F et $(x_1, x_2, \dots, x_n)$ un n -uplet de vecteurs deux à deux distincts de $\mathbb{R}^2$ tel que le déterminant vandermondien $v_{f_1, \dots, f_n}(x_1, \dots, x_n)$ n'est pas nul.

On considère un vecteur y de $\mathbb{R}^2$ non colinéaire au vecteur $x_1 - x_2$ et pour tout réel $r > 0$, on définit l'application γ_r de $[0, 1]$ dans $(\mathbb{R}^2)^n$ par :

$$\forall t \in [0, 1], \gamma_r(t) = ((1-t)x_1 + tx_2 + r \sin(\pi t)y, (1-t)x_2 + tx_1 - r \sin(\pi t)y, x_3, \dots, x_n).$$

a) Soit $r > 0$.

Calculer $\gamma_r(0)$ et $\gamma_r(1)$, puis en déduire qu'il existe un réel t compris entre 0 et 1 pour lequel le déterminant vandermondien $v_{f_1, \dots, f_n}(\gamma_r(t))$ est nul.

b) Justifier que, pour tout $r > 0$ et tout $t \in [0, 1]$, on a :

$$(1-t)x_1 + tx_2 + r \sin(\pi t)y \neq (1-t)x_2 + tx_1 - r \sin(\pi t)y.$$

c) Prouver qu'il existe au plus un élément (t, r) de $]0, 1[\times]0, +\infty[$ pour lequel

$$(1-t)x_1 + tx_2 + r \sin(\pi t)y = x_3.$$

d) En généralisant le résultat précédent, justifier qu'il existe un réel $r > 0$ tel que, pour tout $t \in [0, 1]$, les n vecteurs

$$(1-t)x_1 + tx_2 + r \sin(\pi t)y, (1-t)x_2 + tx_1 - r \sin(\pi t)y, x_3, \dots, x_n$$

de $\mathbb{R}^2$ sont distincts deux à deux.

e) Conclure.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2017

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet se compose d'un exercice et d'un problème dont les différentes parties peuvent, dans une certaine mesure, être traitées de manière indépendante, quitte à admettre les résultats des parties précédentes. La qualité de la rédaction et la précision des réponses apportées entrent pour une part importante dans la notation. Les réponses absurdes sont pénalisées.

Problème

Dans tout le problème, X désigne une variable aléatoire réelle, prenant ses valeurs exclusivement dans l'intervalle $]0, +\infty[$. La fonction de répartition de X est notée F_X . On rappelle sa définition : pour tout $x \in \mathbb{R}$,

$$F_X(x) = \mathbb{P}(X \leq x).$$

On pose en outre deux hypothèses supplémentaires sur X : pour tout $x > 0$, on a $\mathbb{P}(X = x) = 0$, et, d'autre part, pour tous $0 < a < b$, on a $\mathbb{P}(a \leq X \leq b) > 0$. In fine, on a donc les propriétés suivantes pour F_X , qu'il n'est pas nécessaire de redémontrer :

- (i) $F_X(0) = 0$.
- (ii) $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
- (iii) La fonction F_X est strictement croissante sur $]0, +\infty[$.
- (iv) La fonction F_X est continue.

Partie I – Quelques calculs avec la loi uniforme sur $]0, 1[$

Dans cette partie, on considère une variable aléatoire U prenant ses valeurs dans l'intervalle $]0, 1[$, et suivant la loi uniforme sur cet intervalle. La loi de U est donc caractérisée par la fonction de densité f_U définie sur $\mathbb{R}$ par $f_U(u) = 1$ pour $u \in]0, 1[$, et $f_U(u) = 0$ pour $u \notin]0, 1[$.

- 1) On note F_U la fonction de répartition de U . Montrer que, pour tout $u \in [0, 1]$, on a la relation $F_U(u) = u$.
- 2) On considère deux nombres réels $0 \leq a \leq b \leq 1$. Montrer que $\mathbb{P}(a \leq U \leq b) = b - a$.
- 3) Montrer que la variable aléatoire $(1 - U)$ suit également la loi uniforme sur l'intervalle $]0, 1[$.
- 4) Calculer la valeur de $\mathbb{E}(U)$.
- 5) Plus généralement, calculer la valeur de $\mathbb{E}(U^m)$, pour tout entier $m \geq 1$.

Partie II – Fonction quantile q_X

- 6) Montrer que, pour tout $u \in]0, 1[$, il existe un et un seul nombre réel $x \in]0, +\infty[$ vérifiant l'identité $F_X(x) = u$.

Dans toute la suite, on note $q_X(u)$ le nombre réel x dont l'existence est établie à la question 1). On a donc, pour tout $u \in]0, 1[$, l'identité $F_X(q_X(u)) = u$.

- 7) Montrer que, pour tout $u \in]0, 1[$, on a $\mathbb{P}(X \geq q_X(u)) = 1 - u$.
- 8) Montrer que, pour tous $0 < u < v < 1$, on a l'inégalité $q_X(u) < q_X(v)$.
- 9) Montrer que, pour tout $x > 0$, on a équivalence entre la condition $q_X(u) \leq x$ et la condition $u \leq F_X(x)$.
- 10) On considère une variable aléatoire U prenant ses valeurs dans l'intervalle $]0, 1[$, et suivant la loi uniforme sur cet intervalle, comme décrit dans la partie I. Montrer que la variable aléatoire Y définie par $Y = q_X(U)$ suit la même loi de probabilité que la variable aléatoire X .
Indication : calculer la fonction de répartition de Y .
- 11) On définit la variable aléatoire V par $V = F_X(X)$. Montrer que V suit la loi uniforme sur l'intervalle $]0, 1[$.
- 12) Etablir la formule

$$\mathbb{E}(X) = \int_0^1 q_X(u) du.$$

Partie III – Expériences répétées

Dans cette partie, on considère une suite de variables aléatoires $(X_n)_{n \geq 1}$, que l'on suppose indépendantes, et suivant chacune la même loi de probabilité que X . Etant donné un nombre $u \in]0, 1[$, on introduit la variable aléatoire $T_1(u)$ définie par

$$T_1(u) = \inf \{n \geq 1 \mid X_n \geq q_X(u)\},$$

avec la convention $T_1(u) = +\infty$ s'il n'existe aucun entier $n \geq 1$ pour lequel $X_n \geq q_X(u)$. La variable aléatoire $T_1(u)$ correspond donc au plus petit indice n tel que $X_n \geq q_X(u)$.

13) Pour tout entier $k \geq 1$, donner une expression de l'événement $\{T_1(u) = k\}$ en termes des valeurs prises par les variables aléatoires $X_1, \dots, X_k$.

14) Exprimer pour tout entier $k \geq 1$ la valeur de $\mathbb{P}(T_1(u) = k)$.

15) Calculer l'espérance et la variance de $T_1(u)$.

On s'intéresse maintenant à la variable aléatoire $X_{T_1(u)}$ définie par le fait que, pour tout entier $k \geq 1$, $X_{T_1(u)} = X_k$ lorsque $T_1(u) = k$, et (par convention) $X_{T_1(u)} = 0$ lorsque $T_1(u) = +\infty$.

16) Montrer que, pour tout entier $k \geq 1$, et tout $x \geq q_X(u)$, on a

$$\mathbb{P}(X_{T_1(u)} \leq x \mid T_1(u) = k) = \frac{F_X(x)}{1-u} - \frac{u}{1-u}.$$

17) En déduire que, pour tout $x \geq q_X(u)$, on a

$$\mathbb{P}(X_{T_1(u)} \leq x) = \frac{F_X(x)}{1-u} - \frac{u}{1-u}.$$

18) En déduire l'identité

$$\mathbb{E}(X_{T_1(u)}) = q_X(u) + \frac{1}{1-u} \int_{q_X(u)}^{+\infty} (1 - F_X(x)) dx.$$

Indication : On rappelle la formule, valable pour toute variable aléatoire positive W :

$$\mathbb{E}(W) = \int_0^{+\infty} \mathbb{P}(W > w) dw.$$

On introduit maintenant la variable aléatoire $N_n(u)$ définie par

$$N_n(u) = \sum_{j=1}^n \mathbf{1}_{X_j \geq q_X(u)},$$

où la notation $\mathbf{1}_A$ désigne la variable aléatoire indicatrice de l'événement A , c'est-à-dire la variable aléatoire définie par :

$$\mathbf{1}_A = \begin{cases} 1 & \text{si } A \text{ est réalisé} \\ 0 & \text{si } A \text{ n'est pas réalisé} \end{cases}.$$

19) Quelle est la loi de probabilité de la variable aléatoire $N_n(u)$? Donner la valeur de son espérance et de sa variance.

20) Formaliser et justifier, au moyen de la loi des grands nombres, l'affirmation (imprécise) suivante : « sur un grand nombre n de répétitions, la proportion des cas pour lesquels $X_j \geq q_X(u)$ doit être approximativement égale à $1 - u$ ».

21) Etant donné une valeur $u \in]0, 1[$, proposer un moyen d'obtenir une estimation de $q_X(u)$ à partir de la donnée d'une « longue » série de valeurs $X_1, \dots, X_n$.

22) Pour tout $n \geq 1$, et tout $0 < u < 1$, on définit la variable aléatoire $R_n(u)$ par $R_n(u) = 0$ lorsque $N_n(u) = 0$, et, lorsque $N_n(u) > 0$, par

$$R_n(u) = \frac{\sum_{j=1}^n X_j \mathbf{1}_{X_j \geq q_X(u)}}{N_n(u)}.$$

Etudier le comportement de $R_n(u)$ lorsque $n \rightarrow +\infty$. (Justifier votre réponse.)

Partie IV – Deux exemples

On suppose pour commencer que X suit la loi exponentielle de paramètre $\lambda > 0$. La loi de X est donc caractérisée par la fonction de densité f_X définie sur $\mathbb{R}$ par $f_X(x) = \lambda e^{-\lambda x}$ pour $x \geq 0$, et $f_X(x) = 0$ pour $x < 0$.

23) Calculer la valeur de $q_X(u)$ pour tout $u \in]0, 1[$.

24) En déduire que la variable aléatoire $-\frac{1}{\lambda} \ln(U)$ suit la loi exponentielle de paramètre λ , où U est la variable aléatoire étudiée dans la partie I.

25) Calculer la valeur de $\mathbb{E}(X_{T_1(u)})$ en fonction de λ et de u .

26) Comment se comporte le rapport $\frac{\mathbb{E}(X_{T_1(u)})}{q_X(u)}$ lorsque $u \rightarrow 1$?

On suppose à présent que X suit une loi caractérisée par la fonction de densité définie sur $\mathbb{R}$ par

$$f_X(x) = \left(\frac{\alpha}{s}\right) \frac{1}{(1 + x/s)^{\alpha+1}} \text{ pour } x \geq 0, \text{ et } f_X(x) = 0 \text{ pour } x < 0,$$

où $s > 0$ et $\alpha > 1$.

27) Calculer la fonction de répartition de X .

28) Calculer la valeur de $q_X(u)$ pour tout $u \in]0, 1[$.

29) Calculer la valeur de $\mathbb{E}(X_{T_1(u)})$ en fonction de (s, α) et de u .

30) Comment se comporte le rapport $\frac{\mathbb{E}(X_{T_1(u)})}{q_X(u)}$ lorsque $u \rightarrow 1$? Commenter la différence éventuelle avec le résultat obtenu à la question 26).

Partie IV – Avec une loi discrète

Dans cette partie, on considère une variable aléatoire D prenant un nombre fini de valeurs distinctes $0 < x_1 < \dots < x_m$, associées respectivement à des probabilités que l'on note $p_1 = \mathbb{P}(D = x_1), \dots, p_m = \mathbb{P}(D = x_m)$, toutes supposées strictement positives. On pose ensuite $\pi_0 = 0$, et, pour tout $\ell \in \{1, \dots, m\}$, $\pi_\ell = \sum_{j=1}^{\ell} p_j$. Pour tout $u \in]0, 1[$, on définit alors $q_D(u)$ comme le nombre x_ℓ , où $\ell \in \{1, \dots, m\}$ est l'unique entier tel que $\pi_{\ell-1} < u \leq \pi_\ell$.

31) On considère une variable aléatoire U prenant ses valeurs dans l'intervalle $]0, 1[$, et suivant la loi uniforme sur cet intervalle, comme décrit dans la partie I. Montrer que la variable aléatoire G définie par $G = q_D(U)$ suit la même loi de probabilité que la variable aléatoire D .

32) La loi de probabilité de la variable $F_D(D)$ peut-elle être la loi uniforme sur l'intervalle $]0, 1[$?

Exercice

On considère une suite de variables aléatoires $(W_n)_{n \geq 1}$, indépendantes, et suivant chacune la loi normale $\mathcal{N}(0, 1)$. On considère d'autre part une variable aléatoire X_0 suivant la loi normale $\mathcal{N}(0, \sigma_0^2)$, indépendante de la suite $(W_n)_{n \geq 1}$. On définit ensuite par récurrence les variables aléatoires $X_1, X_2, \dots$ via la relation, valable pour tout $n \geq 1$:

$$X_n = \alpha X_{n-1} + \sigma W_n,$$

où α et σ sont deux nombres donnés vérifiant $\sigma > 0$ et $\alpha \in]-1, 1[$.

A) Montrer par récurrence que, pour tout $n \geq 1$, la variable X_n suit une loi normale $\mathcal{N}(0, \sigma_n^2)$, où la suite $(\sigma_n)_{n \geq 1}$ obéit à une relation de récurrence que l'on précisera.

B) On pose $\sigma_\star = \sqrt{\frac{\sigma^2}{1-\alpha^2}}$. Montrer que, si $\sigma_0 = \sigma_\star$, alors pour tout $n \geq 1$, X_n suit la loi normale $\mathcal{N}(0, \sigma_\star^2)$.

C) Montrer que, de manière générale, on a $\lim_{n \rightarrow +\infty} \sigma_n^2 = \sigma_\star^2$.

Dans toute la suite, on suppose que $\sigma_0 = \sigma_\star$.

D) Montrer que, pour tout $n \geq 0$, on a $\text{Cov}(X_n, X_{n+1}) = \alpha \sigma_\star^2$. Les variables X_n et X_{n+1} sont-elles indépendantes en général?

E) Montrer que, pour tout $n \geq 0$ et tout $\ell \geq 1$, on a la relation :

$$X_{n+\ell} = \alpha^\ell X_n + \sum_{k=0}^{\ell-1} \sigma \alpha^k W_{n+\ell-k}.$$

F) En déduire que pour tout $n \geq 0$ et tout $\ell \geq 1$, on a :

$$\text{Cov}(X_n, X_{n+\ell}) = \alpha^\ell \sigma_\star^2.$$

G) Donner une expression de la variance $\mathbb{V}(X_1 + \dots + X_n)$, et en déduire une inégalité de la forme $\mathbb{V}(X_1 + \dots + X_n) \leq Cn$, où C est une constante positive.

H) Déduire de la question précédente que l'on a la convergence en probabilité suivante :

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2018

Épreuve de mathématiques

Durée : 4h

On note $\lfloor x \rfloor$ la partie entière d'un réel x . On rappelle qu'un nombre entier naturel, au moins égal à 2, est dit premier s'il n'est divisible que par 1 et lui-même (donc 1 n'est pas premier).

On note $\mathcal{P} = \{2, 3, 5, 7, 11, \dots\}$ l'ensemble des nombres premiers. On rappelle aussi que tout entier naturel n , au moins égal à 2, se décompose, de façon unique à l'ordre des facteurs près, comme produit de nombres premiers c'est-à-dire qu'il existe $r \in \mathbb{N}^*$, $(p_1, p_2, \dots, p_r) \in \mathcal{P}^r$ et $(\alpha_1, \alpha_2, \dots, \alpha_r) \in (\mathbb{N}^*)^r$ tels que :

$$n = \prod_{k=1}^r p_k^{\alpha_k}.$$

Si a et b sont deux entiers naturels tels que $a \leq b$, la notation $\sum_{\substack{a \leq p \leq b \\ p \in \mathcal{P}}} \alpha_p$ désigne la somme des nombres α_p pour tous les entiers **premiers** p de l'intervalle entier $\llbracket a, b \rrbracket$. On définit de la même façon $\sum_{\substack{p \leq b \\ p \in \mathcal{P}}} \alpha_p$, $\prod_{\substack{a \leq p \leq b \\ p \in \mathcal{P}}} \alpha_p$, etc.

Par exemple, $\sum_{\substack{4 \leq p \leq 10 \\ p \in \mathcal{P}}} \alpha_p = \alpha_5 + \alpha_7$, ou $\prod_{\substack{p \leq 8 \\ p \in \mathcal{P}}} \alpha_p = \alpha_2 \times \alpha_3 \times \alpha_5 \times \alpha_7$.

Partie I - Préliminaires

On établit, dans cette partie, quelques résultats préliminaires, indépendants les uns des autres, qui seront utilisés par la suite.

1. Soit $n_0 \in \mathbb{N}^*$ et f une fonction continue, décroissante et positive de $[n_0, +\infty[$ dans $\mathbb{R}$.

On pose, pour tout entier naturel n non nul, $S_n = \sum_{k=n_0}^n f(k)$.

- (a) Montrer que la suite $(\gamma_n)_{n \geq n_0}$ de terme général $\gamma_n = S_n - \int_{n_0}^n f(t) dt$ est monotone et convergente.

- (b) En déduire, l'existence d'un réel, noté C , pour lesquels on a, lorsque l'entier n tend vers $+\infty$:

$$\sum_{k=2}^n \frac{1}{k \ln k} = \ln \ln n + C + o(1).$$

- (c) Établir la convergence de l'intégrale $\int_2^{+\infty} \frac{1}{t \ln^2 t} dt$ et en déduire la convergence de la série $\sum \frac{1}{k \ln^2 k}$.

2. Montrer que la série de terme général $\frac{\ln k}{k(k-1)}$ est convergente.

On note $K = \sum_{k=2}^{+\infty} \frac{\ln k}{k(k-1)}$ sa somme.

3. (a) Prouver, pour tout entier naturel n au moins égal à 2, l'inégalité :

$$\sum_{k=2}^n \ln k \geq n \ln n - n + 1.$$

- (b) En déduire, quand n tend vers $+\infty$, l'estimation : $\ln n! = n \ln n + O(n)$.

4. (a) Soit λ un réel strictement positif. Justifier, pour tout $n \in \mathbb{N}^*$, l'existence et l'unicité d'un réel $x > 0$ tel que $x \ln x - \lambda x = \ln n$. On note r_n cet unique réel.

- (b) Montrer que $\lim_{n \rightarrow +\infty} r_n = +\infty$ puis établir l'équivalence $r_n \sim \frac{\ln n}{\ln \ln n}$.

5. On note, pour toute partie E de $\mathbb{N}^*$ et pour tout $n \in \mathbb{N}^*$, E_n l'ensemble des éléments de E inférieurs ou égaux à n , c'est-à-dire que $E_n = E \cap \llbracket 1, n \rrbracket$, et l'on pose, pour tout entier naturel n non nul, $d_n(E) = \frac{1}{n} \text{card}(E_n)$.

Si la suite $(d_n(E))_{n \in \mathbb{N}^*}$ converge, on note $d(E)$ sa limite et on dit que la partie E de $\mathbb{N}^*$ admet une densité égale à $d(E)$.

- (a) Montrer que les ensembles suivants possèdent une densité dont on donnera la valeur :

i. Une partie F finie de $\mathbb{N}^*$.

ii. L'ensemble $a\mathbb{N}^* := \{ka; k \in \mathbb{N}^*\}$ des multiples non nuls de l'entier $a \in \mathbb{N}^*$.

- iii. L'ensemble $C := \{k^2; k \in \mathbb{N}^*\}$ des entiers non nuls qui sont des carrés.
- (b) Soit E_1, E_2 deux parties **disjointes** de $\mathbb{N}^*$ possédant une densité. Les parties $\mathbb{N}^* \setminus E_1$ et $E_1 \cup E_2$ possèdent-elles une densité? Et, si oui, que valent-elles?
- (c) L'application d est-elle une probabilité sur l'ensemble $\mathbb{N}^*$ muni de la tribu formée de toutes ses parties?
6. (a) Justifier, pour tout entier naturel m non nul, l'inégalité : $2 \binom{2m+1}{m} \leq 2^{2m+1}$.
- (b) Montrer que, pour tout entier naturel r non nul, l'entier $\prod_{\substack{r+1 < p \leq 2r+1 \\ p \in \mathcal{P}}} p$ divise l'entier $\binom{2r+1}{r}$ (le produit s'effectuant donc sur tous les entiers **premiers** de $\llbracket r+1, 2r+1 \rrbracket$).
- (c) Établir, pour tout entier n au moins égal à 2, l'inégalité : $\prod_{\substack{p \leq n \\ p \in \mathcal{P}}} p \leq 4^n$ (le produit s'effectuant donc sur tous les entiers **premiers** au plus égaux à n).
- On raisonnera par récurrence forte et, ayant supposé l'inégalité vraie jusqu'au rang n , on examinera, en particulier, le cas où $n+1$ est un entier premier égal à $2r+1$.*
- On en déduit ainsi l'inégalité : $\sum_{\substack{p \leq n \\ p \in \mathcal{P}}} \ln p \leq n \ln 4$.
7. Soit $n \in \mathbb{N}^*$. On note, pour tout nombre premier p et tout entier $r \in \mathbb{N}$, $r \geq 2$, $v_p(r)$ l'exposant de p dans la décomposition en nombres premiers de r , et on pose $v_p(1) = 0$. Par exemple, puisque $300 = 2^2 3^2 5^2$, $v_2(300) = 2$, $v_3(300) = 1$, $v_5(300) = 2$ et $v_p(300) = 0$ si $p \notin \{2, 3, 5\}$.
- Soit p un nombre premier. On note, pour tout entier naturel k non nul, α_k (resp. β_k) le nombre d'entiers $d \in \llbracket 1, n \rrbracket$ tels que p^k divise d (resp. tels que $v_p(d) = k$).
- Bien sûr, dès que k est assez grand, $\alpha_k = \beta_k = 0$.*
- (a) Prouver, pour tout $k \in \mathbb{N}^*$, l'égalité : $\alpha_k = \left\lfloor \frac{n}{p^k} \right\rfloor$.
- (b) Justifier l'égalité : $v_p(n!) = \sum_{k=1}^{+\infty} k \beta_k$.
- (c) En déduire, en reliant β_k aux α_i , l'égalité : $v_p(n!) = \sum_{k=1}^{+\infty} \left\lfloor \frac{n}{p^k} \right\rfloor$.
- (d) En déduire l'encadrement : $\frac{n}{p} - 1 \leq v_p(n!) \leq \frac{n}{p-1}$ ($= \frac{n}{p} + \frac{n}{p(p-1)}$).
8. Soit $(a_n)_{n \in \mathbb{N}^*}$ et $(\varepsilon_n)_{n \in \mathbb{N}^*}$ deux suites réelles. On note, pour tout $n \in \mathbb{N}^*$, $A_n = \sum_{k=1}^n a_k$.
- Prouver, pour tout entier $n \geq 2$, l'égalité

$$\sum_{k=1}^n \varepsilon_k a_k = \sum_{k=1}^{n-1} (\varepsilon_k - \varepsilon_{k+1}) A_k + \varepsilon_n A_n.$$

9. Soit $(a_N)_{N \in \mathbb{N}^*}$ une suite positive de limite $+\infty$ et $(b_N)_{N \in \mathbb{N}^*}$ une suite bornée. Soit $(X_N)_{N \in \mathbb{N}^*}$ une suite de variables aléatoires discrètes, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$. On suppose que, pour tout $N \in \mathbb{N}^*$, $\mathbf{E}(X_N) = a_N + b_N$ et, quand N tend vers $+\infty$, $\mathbf{Var}(X_N) = O(a_N)$.

(a) Justifier, pour tout entier N assez grand, l'inclusion entre événements :

$$\left[|X_N - \mathbf{E}(X_N)| \leq \frac{1}{2} a_N^{2/3} \right] \subset \left[|X_N - a_N| \leq a_N^{2/3} \right].$$

(b) En déduire que

$$\lim_{N \rightarrow +\infty} \mathbf{P}(|X_N - a_N| > a_N^{2/3}) = 0.$$

Partie II - Deux résultats asymptotiques

1. (a) Établir, pour tout entier naturel n non nul, l'égalité : $\ln n! = \sum_{\substack{p \leq n \\ p \in \mathcal{P}}} \nu_p(n!) \ln p$.

(b) En déduire, pour tout entier naturel n non nul, l'encadrement :

$$\frac{\ln n!}{n} - K \leq \sum_{\substack{p \leq n \\ p \in \mathcal{P}}} \frac{\ln p}{p} \leq \frac{\ln n!}{n} + \ln 4,$$

où le réel K est défini dans la question I - 2).

(c) Conclure, quand l'entier n tend vers $+\infty$, à l'évaluation asymptotique

$$\sum_{\substack{p \leq n \\ p \in \mathcal{P}}} \frac{\ln p}{p} = \ln n + O(1).$$

2. On note χ l'application qui, à chaque entier $k \in \mathbb{N}^*$, associe 1 si k est premier (i.e. $k \in \mathcal{P}$) et 0 sinon.

(a) En posant, pour tout pour tout entier naturel k non nul, $a_k = \chi(k) \frac{\ln k}{k}$, $A_k = \sum_{i=1}^k a_i$, en utilisant I - 8), établir, pour tout $n \geq 2$, l'égalité :

$$\sum_{\substack{p \leq n \\ p \in \mathcal{P}}} \frac{1}{p} = \sum_{k=2}^{n-1} \frac{\ln(1+1/k)}{\ln k \ln(k+1)} A_k + \frac{A_n}{\ln n}.$$

(b) Établir, quand l'entier k tend vers $+\infty$, l'égalité : $\frac{\ln(1+1/k)}{\ln k \ln(k+1)} A_k = \frac{1}{k \ln k} + O\left(\frac{1}{k \ln^2 k}\right)$.

(c) En déduire, quand l'entier n tend vers $+\infty$, l'égalité :

$$\sum_{\substack{p \leq n \\ p \in \mathcal{P}}} \frac{1}{p} = \ln \ln n + O(1).$$

Partie III

On note, pour tout entier naturel n non nul, $\omega(n)$ le nombre d'entiers **premiers** distincts qui divisent l'entier n . On a donc $\omega(2^5) = 1$, $\omega(2^2 \cdot 5^3) = 2$, $\omega(2 \cdot 5^2 \cdot 11^5) = 3$, etc.

L'objet de la suite du problème est le contrôle asymptotique, en un certain sens, de la suite $(\omega(n))_{n \in \mathbb{N}^*}$.

1. (a) Soit $n \in \mathbb{N}$, $n \geq 2$, dont la décomposition en nombres premiers est $n = \prod_{k=1}^r p_k^{\alpha_k}$.

On a donc $\omega(n) = r$.

Prouver l'inégalité : $\omega(n) = r \leq \frac{\ln n}{\ln 2}$.

- (b) À l'aide de I-4) prouver la domination $\omega(n) = O\left(\frac{\ln n}{\ln \ln n}\right)$.

On observera que $n \geq 2 \prod_{k=1}^{r-1} (2k+1)$ puis on prouvera, pour un réel λ qu'on déterminera, l'inégalité : $\ln n \geq (r-1) \ln(r-1) - \lambda(r-1)$.

2. Soit $N \in \mathbb{N}^*$. On munit l'ensemble $\llbracket 1, N \rrbracket$ de la probabilité uniforme, et, pour tout $r \in \llbracket 1, N \rrbracket$, on note $X_{N,r}$ la variable aléatoire suivante :

$$X_{N,r} : \begin{array}{ccc} \llbracket 1, N \rrbracket & \longrightarrow & \mathbb{R} \\ d & \longmapsto & \begin{cases} 1 & \text{si } r \text{ divise } d \\ 0 & \text{sinon} \end{cases}, \end{array}$$

et note $X_N = \sum_{\substack{p \leq N \\ p \in \mathcal{P}}} X_{N,p}$ (on effectue donc la somme sur tous les entiers p **premiers** de $\llbracket 1, N \rrbracket$).

On a donc, pour tout $n \in \llbracket 1, N \rrbracket$, $X_N(n) = \omega(n)$.

On note $\mathbf{E}(Y)$ et $\mathbf{Var}(Y)$ l'espérance et la variance d'une variable aléatoire Y sur l'espace précédent (elles dépendent, bien sûr, de l'entier N).

- (a) Vérifier, pour tout $r \in \llbracket 1, N \rrbracket$, l'égalité : $\mathbf{E}(X_{N,r}) = \frac{1}{N} \left\lfloor \frac{N}{r} \right\rfloor$.
- (b) Prouver l'égalité : $\mathbf{E}(X_N^2) = \mathbf{E}(X_N) + \sum_{\substack{1 \leq p, q \leq N \\ (p,q) \in \mathcal{P}^2 \text{ et } p \neq q}} \frac{1}{N} \left\lfloor \frac{N}{pq} \right\rfloor$.
- (c) En déduire, quand l'entier N tend vers $+\infty$, l'ordre de grandeur : $\mathbf{Var}(X_N) = O(\ln \ln N)$.
- (d) En déduire, à l'aide d'un résultat de la **partie I**, le résultat :

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \text{card} \left\{ n \in \llbracket 1, N \rrbracket ; |\omega(n) - \ln \ln N| > (\ln \ln N)^{2/3} \right\} = 0.$$

Ainsi, en examinant le cas où $N = 10^{99}$, on peut s'attendre à ce que, « le plus souvent », un entier d'au plus 100 chiffres, possède entre 3 et 8 facteurs premiers distincts. Étonnant non ?

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2018

Épreuve à option (A) : Mathématiques

Durée : 4h

Dans tout l'énoncé, p désigne un nombre entier strictement positif et $(e_1, \dots, e_p)$ la base canonique de $\mathbb{R}^p$.

Pour toute matrice carrée $A \in \mathcal{M}_p(\mathbb{R})$, on appelle *solution de l'équation de réplication de matrice A*, toute application $f : t \mapsto (f_1(t), f_2(t), \dots, f_p(t))$ de $\mathbb{R}$ dans $\mathbb{R}^p$, dont les composantes $f_1, f_2, \dots, f_p$ sont des fonctions dérivables sur $\mathbb{R}$ qui vérifient :

$$\boxed{\forall i \in \llbracket 1, p \rrbracket, \forall t \in \mathbb{R}, \quad f'_i(t) = \langle e_i - f(t), Af(t) \rangle f_i(t)} \quad (1)$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire usuel sur $\mathbb{R}^p$ et $Af(t)$ le vecteur de $\mathbb{R}^p$ dont la matrice-

colonne dans la base canonique est la matrice-produit $A \times \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_p(t) \end{pmatrix}$.

L'objet du problème est d'étudier la trajectoire des solutions de (1) et en particulier leurs limites en $+\infty$.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Partie I : solutions de l'équation de réplication scalaire

Dans le cas où p est égal à 1, les matrices de $\mathcal{M}_p(\mathbb{R})$ ne possèdent qu'un seul coefficient et l'équation (1) se réduit à une équation différentielle scalaire, dont l'étude fait l'objet de cette partie.

Soit a un nombre réel différent de 0.

On se propose de montrer que, pour tout réel $y \in]0, 1[$, il existe une unique fonction x définie et de classe C^1 sur $\mathbb{R}$, à valeurs dans $]0, 1[$ vérifiant :

$$\begin{cases} x(0) = y \\ \forall t \in \mathbb{R}, \quad x'(t) = a(x(t))^2(1 - x(t)) \end{cases} \quad (2)$$

1. On note φ l'application définie sur $]0, 1[$ par :

$$\forall x \in \mathbb{R}, \quad \varphi(x) = -\frac{1}{x} + \ln\left(\frac{x}{1-x}\right) \quad (3)$$

- a) Justifier que φ est une bijection de $]0, 1[$ sur $\mathbb{R}$.
 - b) Donner l'allure du graphe de φ .
 - c) Les fonctions φ et φ^{-1} sont-elles lipschitziennes?
2. a) Pour tout $t \in \mathbb{R}$ et toute fonction f de classe C^1 sur $\mathbb{R}$, à valeurs dans $]0, 1[$, exprimer l'intégrale $\int_0^t \frac{f'(u)}{(f(u))^2(1-f(u))} du$ à l'aide des fonctions φ et f , de t et de $f(0)$.

b) Démontrer que, pour tout $y \in \mathbb{R}$, l'unique fonction x définie et de classe C^1 sur $\mathbb{R}$, à valeurs dans $]0, 1[$, qui vérifie (2) est la fonction f_y donnée par :

$$\forall t \in \mathbb{R}, \quad f_y(t) = \varphi^{-1}(a t + \varphi(y)).$$

3. On note $\mathcal{B}(\mathbb{R}, \mathbb{R})$ le $\mathbb{R}$ -espace vectoriel des fonctions bornées à valeurs dans $\mathbb{R}$, muni de la norme de la convergence uniforme.

a) Démontrer que l'application Φ qui associe à tout élément y de $]0, 1[$ la fonction f_y est une application continue de $\mathbb{R}$ dans l'espace vectoriel normé $\mathcal{B}(\mathbb{R}, \mathbb{R})$.

b) On note $\mathcal{S} = \{f_y; y \in]0, 1[\}$.

Justifier que $\mathcal{S}$ est une partie connexe par arcs de l'espace vectoriel normé $\mathcal{B}(\mathbb{R}, \mathbb{R})$.
 $\mathcal{S}$ est-elle une partie ouverte de $\mathcal{B}(\mathbb{R}, \mathbb{R})$? En est-elle une partie fermée?

Partie II : étude du cas où $p = 2$

Dans cette partie, on suppose que p est égal à 2 et on note $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$.

Soit $x_0 \in]0, 1[$.

On admet qu'il existe une unique application $f : t \mapsto (f_1(t), f_2(t))$ de $\mathbb{R}$ dans $\mathbb{R}^2$ vérifiant (1) et telle que :

$$\begin{cases} f_1(0) = x_0 \\ f_2(0) = 1 - x_0 \end{cases} \quad (4)$$

1. Soit g une fonction définie et dérivable sur $\mathbb{R}$, telle que g et sa dérivée g' admettent chacune une limite finie en $+\infty$.

Démontrer que la limite de la dérivée g' en $+\infty$ est nécessairement nulle.

2. Soit h une fonction continue sur $\mathbb{R}$, à valeurs réelles et H une primitive de h sur $\mathbb{R}$.
On considère une fonction $x : t \mapsto x(t)$ définie et dérivable sur $\mathbb{R}$ vérifiant :

$$\forall t \in \mathbb{R}, x'(t) = h(t)x(t).$$

a) Donner, pour tout $t \in \mathbb{R}$, une expression de $x(t)$ en fonction de $x(0)$ et deux valeurs de la fonction H .

b) Que peut-on en déduire sur le signe de la fonction x si $x(0)$ n'est pas nul?

3. a) Justifier, pour tout réel t , l'égalité :

$$f_1'(t) + f_2'(t) = \langle f(t), Af(t) \rangle (1 - f_1(t) - f_2(t)).$$

b) Justifier que $f_1(t) + f_2(t)$ est égal à 1 pour tout réel t .

4. On suppose dans cette question que c est égal à d .

a) Utiliser les résultats de la partie I pour exprimer f_1 à l'aide de la fonction φ .

b) En déduire la limite de $f(t)$ quand t tend vers $+\infty$ selon les valeurs de a et b .

5. On suppose dans cette question que a et d sont égaux et non nuls, et que b et c sont nuls, autrement dit

$$A = a I_2 \quad \text{avec } a \neq 0.$$

a) Justifier que, pour tout $t \in \mathbb{R}$, on a :

$$f_1'(t) = a f_1(t) (1 - f_1(t)) (2f_1(t) - 1).$$

b) On suppose dans cette sous-question que x_0 est strictement supérieur à $\frac{1}{2}$.

i) Justifier que, pour tout $t \in \mathbb{R}$, on a : $\frac{1}{2} < f_1(t) < 1$.

ii) En déduire que $f_1(t)$ admet une limite quand t tend vers $+\infty$.

iii) Trouver, selon le signe de a , la limite de $f(t)$ quand t tend vers $+\infty$.

c) Étudier la convergence de $f(t)$ quand t tend vers $+\infty$ dans le cas où $x_0 < \frac{1}{2}$.

Partie III : inégalités de Pinsker

1. On considère la fonction K définie sur l'ouvert $]0, 1[\times]0, 1[$ de $\mathbb{R}^2$ par :

$$\forall (x, y) \in]0, 1[\times]0, 1[, \quad K(x, y) = x \ln\left(\frac{x}{y}\right) + (1-x) \ln\left(\frac{1-x}{1-y}\right) \quad (5)$$

a) En utilisant la concavité de la fonction $\ln$, démontrer que la fonction K est minorée. Est-elle majorée?

- b) Justifier que K est de classe C^1 sur $]0, 1[\times]0, 1[$ et calculer les dérivées partielles $\partial_1 K$ et $\partial_2 K$.
- c) En déduire les points où la fonction K atteint sa borne inférieure.
- d) Justifier l'inégalité :

$$x \ln\left(\frac{x}{y}\right) + (1-x) \ln\left(\frac{1-x}{1-y}\right) \geq 2(x-y)^2 \quad (6)$$

2. Pour toute partie D de $\llbracket 1, p \rrbracket$ et tout vecteur $x = (x_1, x_2, \dots, x_p)$ de $\mathbb{R}^p$, on note :

$$x_D = \begin{cases} \sum_{i \in D} x_i & \text{si } D \neq \emptyset \\ 0 & \text{si } D = \emptyset \end{cases}$$

Soit $x = (x_1, x_2, \dots, x_p)$ et $y = (y_1, y_2, \dots, y_p)$ deux éléments de $\mathbb{R}_+^p$ tels que :

$$\sum_{i=1}^p x_i = \sum_{i=1}^p y_i = 1.$$

On note $B_+ = \{i \in \llbracket 1, p \rrbracket \mid x_i > y_i\}$ et $B_- = \{i \in \llbracket 1, p \rrbracket \mid x_i \leq y_i\}$.

a) Exprimer $\sum_{i=1}^p |x_i - y_i|$ en fonction de x_{B_+} et y_{B_+} .

b) Justifier l'inégalité :

$$\sum_{i=1}^p x_i \ln\left(\frac{x_i}{y_i}\right) \geq x_{B_+} \ln\left(\frac{x_{B_+}}{y_{B_+}}\right) + (1 - x_{B_+}) \ln\left(\frac{1 - x_{B_+}}{1 - y_{B_+}}\right).$$

c) En déduire l'inégalité, dite de Pinsker, qui généralise l'inégalité (6) :

$$\sum_{i=1}^p x_i \ln\left(\frac{x_i}{y_i}\right) \geq \frac{1}{2} \left(\sum_{i=1}^p |x_i - y_i| \right)^2 \quad (7)$$

Partie IV : Convergence vers un point de coordonnées strictement positives

On note $\Delta = \{(x_1, \dots, x_p) \in \mathbb{R}_+^p \mid \sum_{i=1}^p x_i = 1\}$ et $\Delta^0 = \{(x_1, \dots, x_p) \in \mathbb{R}_+^p \mid \sum_{i=1}^p x_i = 1\}$.

Soit f une solution de l'équation de réplication (1) telle que $f(0)$ appartient à Δ^0 , c'est-à-dire

$$\begin{cases} \forall i \in \llbracket 1, p \rrbracket, f_i(0) > 0 \\ f_1(0) + f_2(0) + \dots + f_p(0) = 1 \end{cases}.$$

1. Justifier les deux assertions :

a) $\forall t \in \mathbb{R}, \sum_{i=1}^p f_i(t) = 1$.

b) f est à valeurs dans Δ^0 .

On suppose désormais qu'il existe un vecteur $x^* = (x_1^*, x_2^*, \dots, x_p^*) \in \Delta^0$ tel que :

$$\forall x \in \Delta \setminus \{x^*\}, \langle x^* - x, Ax \rangle > 0 \quad (8)$$

2. On note Q la fonction définie sur l'ouvert $]0, 1[^p$ de $\mathbb{R}^p$ par :

$$\forall x = (x_1, x_2, \dots, x_p) \in]0, 1[^p, \quad Q(x) = \sum_{i=1}^p x_i^* \ln\left(\frac{x_i^*}{x_i}\right) \quad (9)$$

- a) Justifier que $Q(x)$ est positif ou nul pour tout $x \in \Delta^0$.
- b) Justifier que x^* est l'unique élément x de Δ^0 tel que $Q(x) = 0$.
- c) Pour tout $x \in \Delta^0$, justifier les inégalités :

$$Q(x) \leq \sum_{i=1}^p \frac{x_i^*}{x_i} (x_i^* - x_i) \leq \frac{1}{\min\{x_1, x_2, \dots, x_p\}} \left(\sum_{i=1}^p (x_i - x_i^*)^2 \right)^{1/2}.$$

- 3. a) Justifier que la fonction composée $Q \circ f$ est de classe C^1 et exprimer sa dérivée à l'aide de f , A et x^* .
- b) En déduire que $Q \circ f$ admet une limite positive ou nulle ℓ en $+\infty$.

4. On suppose dans cette question qu'il existe un réel strictement positif ε tel que :

$$\forall t \in \mathbb{R}_+, \forall i \in \llbracket 1, p \rrbracket, f_i(t) \geq \varepsilon \quad (10)$$

- a) Justifier, pour tout $t \in \mathbb{R}_+$, l'inégalité : $\sum_{i=1}^p (f_i(t) - x_i^*)^2 \geq \varepsilon^2 \ell^2$.
- b) Justifier que, pour tout réel strictement positif α , il existe $\beta > 0$ tel que :

$$\forall x \in \Delta, \quad \left(\langle x - x^*, x - x^* \rangle \geq \alpha \right) \implies \left(\langle x^* - x, Ax \rangle \geq \beta \right).$$

- c) En déduire, en raisonnant par l'absurde, que la limite ℓ de $Q \circ f$ en $+\infty$ est nulle.
- d) Démontrer que $f(t)$ tend vers x^* quand t tend vers $+\infty$.

5. Un exemple

Dans cette question, on suppose $p = 3$ et $A = \begin{pmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & -1 & -1 \end{pmatrix}$.

- a) Justifier l'existence d'un unique vecteur x^* vérifiant (8) et le trouver.
- b) Démontrer que la fonction $t \mapsto f_1(t)f_2(t)f_3(t)$ est croissante.
- c) Justifier que $f(t)$ tend vers x^* quand t tend vers $+\infty$.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2018

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet se compose d'un exercice et d'un problème dont les différentes parties peuvent, dans une certaine mesure, être traitées de manière indépendante, quitte à admettre les résultats des parties précédentes. La qualité de la rédaction et la précision des réponses apportées entrent pour une part importante dans la notation. Les réponses absurdes sont pénalisées.

Problème

Dans tout le problème, n désigne un nombre entier, et l'on suppose que $n \geq 2$. On considère n variables aléatoires notées $X_1, \dots, X_n$, avec les hypothèses suivantes :

(α) Pour tout $1 \leq k \leq n$, la variable X_k ne peut prendre comme valeur que 0 ou 1.

(β) Pour tout $1 \leq k \leq n$, on a $0 < \mathbb{P}(X_k = 1) < 1$.

(γ) Les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes.

Dans la suite, on notera, pour tout $1 \leq k \leq n$:

$$p_k = \mathbb{P}(X_k = 1), \quad q_k = 1 - p_k, \quad r_k = \frac{p_k}{q_k}.$$

Etant donné deux entiers naturels a et b , on utilisera dans la suite la notation $\llbracket a, b \rrbracket$ pour désigner l'ensemble des entiers j tels que $a \leq j \leq b$ (ainsi $\llbracket a, b \rrbracket$ correspond à l'ensemble vide si l'on a $a > b$).

Partie I – Calculs avec les fonctions génératrices

Pour tout nombre réel $t \in [0, +\infty[$, et tout nombre entier $j \in \llbracket 1, n \rrbracket$, on définit

$$g_j(t) = \mathbb{E}(t^{X_j}),$$

en posant par convention que $0^0 = 1$. On définit également, pour $j \in \llbracket 0, n-1 \rrbracket$ la variable aléatoire S_j par

$$S_j = \sum_{k=j+1}^n X_k. \quad (1)$$

On pose en outre

$$G_j(t) = \mathbb{E}(t^{S_j}),$$

toujours avec la convention que $0^0 = 1$.

1) Montrer que l'on a, pour tout $t \in [0, +\infty[$, et $j \in \llbracket 1, n \rrbracket$, l'identité :

$$g_j(t) = q_j + p_j t. \quad (2)$$

2) En déduire que l'on a, pour tout $t \in [0, +\infty[$, et $j \in \llbracket 0, n-1 \rrbracket$, l'identité :

$$G_j(t) = \prod_{k=j+1}^n q_k (1 + r_k t). \quad (3)$$

3) A partir de (3), obtenir une expression explicite pour la dérivée $G'_j(t)$ et montrer que

$$G'_j(0) = \left(\prod_{k=j+1}^n q_k \right) \left(\sum_{k=j+1}^n r_k \right). \quad (4)$$

4) Montrer que l'on a l'identité :

$$\mathbb{P}(S_j = 1) = G'_j(0). \quad (5)$$

5) Donner une expression de l'espérance $\mathbb{E}(S_j)$ (à partir de l'expression de G_j , ou en revenant à la définition de S_j).

6) Donner une expression de la variance $\mathbb{V}(S_j)$ (à partir de l'expression de G_j , ou en revenant à la définition de S_j).

Partie II – Probabilité de s'arrêter au dernier succès

Dans cette partie, on considère un entier $s \in \llbracket 0, n-1 \rrbracket$, et l'on note T_s la variable aléatoire définie de la manière suivante :

$$T_s = \min \{j \in \llbracket s+1, n \rrbracket \mid X_j = 1\},$$

avec par convention $T_s = n$ s'il n'existe aucun entier $j \in \llbracket s+1, n \rrbracket$ pour lequel $X_j = 1$.

Lorsque les variables aléatoires $X_1, \dots, X_n$ représentent les résultats (1 = succès, 0 = échec) de n expériences successives, la variable aléatoire T_s correspond donc à l'indice du premier succès strictement postérieur à s lorsqu'un tel succès existe, et prend la valeur n sinon.

On introduit maintenant l'événement A_s défini par :

$$A_s = \{X_{T_s} = 1 \text{ et } X_j = 0 \text{ pour tout } j \in \llbracket T_s + 1, n \rrbracket\}.$$

L'événement A_s correspond donc au fait qu'il existe un succès strictement postérieur à s , celui-ci étant de plus le dernier succès de la séquence $X_1, \dots, X_n$.

7) Petit exemple : on suppose que $n = 5$ et que les valeurs effectivement réalisées des variables aléatoires $X_1, \dots, X_5$ sont les suivantes : $X_1 = 0, X_2 = 1, X_3 = 1, X_4 = 0, X_5 = 0$. Dans cette situation, quelles valeurs les variables aléatoires $T_0, \dots, T_4$ prennent-elles ? Parmi les événements $A_0, \dots, A_4$, quels sont ceux qui sont réalisés ?

8) Montrer que, pour tout $s \in \llbracket 0, n-1 \rrbracket$, on a l'identité $\mathbb{P}(A_s) = \mathbb{P}(S_s = 1)$.

9) On considère un entier $s \in \llbracket 0, n-2 \rrbracket$. Montrer que l'on a l'inégalité $\mathbb{P}(A_s) \leq \mathbb{P}(A_{s+1})$ si et seulement si l'on a l'inégalité $\sum_{k=s+2}^n r_k \geq 1$.

Indication : Faire appel aux questions précédentes, en notant aussi que $\left(\frac{q_{s+1}}{1-q_{s+1}}\right) r_{s+1} = 1$.

10) On suppose que l'on a $\sum_{k=2}^n r_k \geq 1$, et l'on pose

$$s_n^* = \max \left\{ s \in \llbracket 1, n-1 \rrbracket \mid \sum_{k=s+1}^n r_k \geq 1 \right\}.$$

(a) Montrer que s_n^* est bien défini (comme max. d'un ensemble d'entiers fini non-vidé).

(b) Montrer que la suite $(u_m)_{m=1, \dots, n}$ définie par $u_m = \sum_{k=m}^n r_k$ est décroissante.

(c) Montrer que, pour tout $s \in \llbracket 1, s_n^* \rrbracket$, on a $\sum_{k=s+1}^n r_k \geq 1$.

(d) Montrer que, pour tout $s \in \llbracket 0, s_n^* - 1 \rrbracket$, on a $\mathbb{P}(A_s) \leq \mathbb{P}(A_{s+1})$.

(e) Montrer que, pour tout $s \in \llbracket s_n^* + 1, n-1 \rrbracket$, on a $\sum_{k=s+1}^n r_k < 1$.

(f) Montrer que, pour tout $s \in \llbracket s_n^*, n-2 \rrbracket$ on a $\mathbb{P}(A_s) \geq \mathbb{P}(A_{s+1})$.

(g) En déduire que la probabilité $\mathbb{P}(A_s)$ est maximale pour $s = s_n^*$, parmi les différentes valeurs de $s \in \llbracket 0, n-1 \rrbracket$ possibles.

11) On suppose maintenant que l'on a $\sum_{k=2}^n r_k < 1$. Montrer que la probabilité $\mathbb{P}(A_s)$ est maximale pour $s = 0$.

12) On suppose que toutes les probabilités p_i sont égales : il existe un nombre $p \in]0, 1[$ telle que $p_1 = \dots = p_n = p$. Exprimer la valeur de s_n^* en fonction de n et p (en distinguant éventuellement différents cas).

13) Donner une expression de $\mathbb{E}(T_s | A_r)$ et de $\mathbb{E}(T_s | A_r^c)$ en fonction de n et des paramètres p_j, q_j, r_j . (On note A_r^c l'événement complémentaire de A_r .)

Partie III – Records d'une permutation aléatoire

On rappelle qu'une permutation des entiers de 1 à n est une bijection de l'ensemble $\{1, \dots, n\}$ sur lui-même, autrement dit, une application $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ telle que, pour tout $j \in \llbracket 1, n \rrbracket$, il existe un et un seul $i \in \llbracket 1, n \rrbracket$ tel que $\sigma(i) = j$. Dans cette partie, on note $\mathcal{S}_n$ l'ensemble des permutations des entiers de 1 à n , et l'on rappelle que cet ensemble comporte $n!$ éléments. On considère ensuite une permutation aléatoire Θ uniformément distribuée sur $\mathcal{S}_n$, autrement dit, Θ est une variable aléatoire à valeurs dans $\mathcal{S}_n$ vérifiant, pour tout $\sigma \in \mathcal{S}_n$, l'identité $\mathbb{P}(\Theta = \sigma) = \frac{1}{n!}$.

Etant donné un entier $1 \leq i \leq n$, on introduit l'événement R_i défini par

$$R_i = \{\Theta(i) > \Theta(j) \text{ pour tout } j \in \llbracket 1, i-1 \rrbracket\}.$$

On définit ensuite Y_i comme la variable aléatoire indicatrice de l'événement R_i , autrement dit : $Y_i = 1$ lorsque R_i se réalise, et $Y_i = 0$ dans le cas contraire.

On admettra sans preuve les deux propriétés (remarquables!) suivantes :

- (a) Pour tout $i \in \llbracket 1, n \rrbracket$, on a $\mathbb{P}(Y_i = 1) = 1/i$.
- (b) Les variables aléatoires $Y_1, \dots, Y_n$ sont mutuellement indépendantes.

On pose dans la suite de cette partie $X_i = Y_i$ pour tout $i \in \llbracket 1, n \rrbracket$.

14) L'une des hypothèses effectuées au début du problème n'est pas vérifiée, avec la définition des $(X_i)_{1 \leq i \leq n}$ posée ci-dessus. Laquelle?

On admettra que, malgré l'hypothèse non-vérifiée identifiée à la question précédente, le résultat de la question 10) demeure néanmoins valable.

15) Montrer que l'événement A_r coïncide avec l'événement $\{\Theta(T_r) = n\}$.

16) Exprimer la valeur de $\sum_{k=2}^n r_k$ dans le cas étudié dans cette partie, et justifier que l'on se trouve dans le cas de la question 10).

17) Montrer que, pour tout couple d'entiers a et b tels que $2 \leq a \leq b$, on a :

$$\ln\left(\frac{b+1}{a}\right) \leq \sum_{\ell=a}^b 1/\ell \leq \ln\left(\frac{b}{a-1}\right).$$

Indication : Comparer la somme $\sum_{\ell=a}^b 1/\ell$ avec l'intégrale de la fonction $x \mapsto 1/x$ sur des intervalles appropriés.

18) En déduire que

$$\lim_{n \rightarrow +\infty} \frac{s_n^*}{n} = 1/e. \quad (6)$$

Indication : Utiliser la définition de s_n^* afin d'obtenir deux inégalités pour lesquelles on peut utiliser la question 17).

19) Voici un contexte d'application possible pour les résultats précédents¹. On dispose d'une liste de n candidatures sur un poste à pourvoir. Les candidat(e)s passent un entretien individuel visant à mesurer leur aptitude vis-à-vis du poste, mais la contrainte suivante est imposée : à l'issue de chaque entretien individuel, une réponse immédiate (recrutement ou refus) doit être donnée à la personne ayant passé l'entretien, si bien que le recrutement s'achève

1. Ou, du moins, une motivation pratique justifiant l'étude de ce type de modèle. Il ne s'agit pas d'une application réaliste.

dès qu'une réponse positive a été donnée à un(e) candidat(e). On suppose que le nombre n de candidatures disponibles est élevé, et la procédure retenue pour mener le recrutement est la suivante : dans une première phase, on fait passer un entretien à un nombre s (fixé à l'avance) de candidat(e)s, qui ne recevront de toute façon pas de réponse positive. Dans une seconde phase, on procède à des entretiens avec les autres candidat(e)s, en donnant une réponse positive dès le premier entretien montrant une aptitude supérieure à l'ensemble des entretiens qui l'ont précédé (ce qui inclut les s entretiens de la première phase, plus les précédents entretiens de la seconde phase), et l'on arrête alors de faire passer des entretiens. Enfin, dans le cas où les entretiens ont continué jusqu'à la n -ème candidature, celle-ci est de toute façon retenue.

- Quels peuvent être les inconvénients liés à un choix de s trop faible ?
- Quels peuvent être les inconvénients liés à un choix de s trop élevé ?
- En supposant que les aptitudes successives des candidat(e)s soient ordonnées selon une permutation aléatoire uniformément distribuée, que suggèrent les résultats précédents quant au choix du s permettant de maximiser la probabilité de recruter la candidature présentant la plus grande aptitude parmi les n disponibles ?
- Quelles critiques formulerez-vous vis-à-vis de la procédure de recrutement envisagée ? Quelle autre procédure (respectant la contrainte d'une réponse immédiate après l'entretien) proposeriez-vous ?
- Quelles critiques formulerez-vous vis-à-vis du modèle probabiliste adopté ?

Exercice

La loi de Gumbel est caractérisée par la fonction de répartition F suivante : pour tout $x \in \mathbb{R}$, $F(x) = e^{-e^{-x}}$.

A) Vérifier que la formule ci-dessus possède bien les propriétés qui caractérisent une fonction de répartition continue :

- $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$;
- F est croissante ;
- F est continue.

B) Montrer que la loi de Gumbel est associée à une fonction de densité, dont on donnera l'expression.

C) Donner un équivalent simple de $1 - F(x)$ lorsque $x \rightarrow +\infty$.

D) Montrer que, si U suit la loi uniforme sur $]0, 1[$, la variable aléatoire définie par $G = -\ln(-\ln(U))$ suit la loi de Gumbel. On rappelle la densité de la loi uniforme sur $]0, 1[$: $f(x) = 0$ pour $x \notin]0, 1[$, et $f(x) = 1$ pour $x \in]0, 1[$.

E) Etant donné un entier $N \geq 1$, on considère N variables aléatoires i.i.d. $X_1, \dots, X_N$ suivant chacune la loi de Gumbel. Montrer que la variable aléatoire Z définie par

$$Z = \max(X_1, \dots, X_N) - \ln N$$

suit également la loi de Gumbel.

Indication : Passer par la fonction de répartition de Z .

F) On considère une suite de variables aléatoires i.i.d. $(Y_n)_{n \geq 1}$ suivant chacune la loi exponentielle $\mathcal{E}(1)$. Montrer que l'on a, pour tout $x \geq 0$, la convergence suivante :

$$\lim_{N \rightarrow +\infty} \mathbb{P}(\max(Y_1, \dots, Y_N) - \ln N \leq x) = e^{-e^{-x}}.$$

On rappelle la densité de la loi exponentielle $\mathcal{E}(1)$: $f_{\mathcal{E}(1)}(x) = 0$ pour $x < 0$, et $f_{\mathcal{E}(1)}(x) = e^{-x}$ pour $x \geq 0$.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2019

Épreuve de mathématiques

Durée : 4h

Toutes les variables aléatoires qui interviennent dans ce problème sont discrètes et définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$.

On note $\lfloor x \rfloor$ la partie entière d'un réel x . On rappelle que l'indicatrice $\mathbf{1}_A$ d'un événement A est la variable aléatoire qui, à chaque $\omega \in \Omega$, associe $\mathbf{1}_A(\omega) = 1$ si $\omega \in A$ et $\mathbf{1}_A(\omega) = 0$ sinon.

On rappelle qu'une fonction réelle h , définie sur $\mathbb{R}$, est dite convexe si, pour tout couple $(x, y) \in \mathbb{R}^2$ et tout réel $\lambda \in [0, 1]$, on a : $h(\lambda x + (1 - \lambda)y) \leq \lambda h(x) + (1 - \lambda)h(y)$. Le candidat se rappellera, ou admettra, que si une fonction h est définie sur $\mathbb{R}$ et convexe alors, pour tout réel a , la fonction $p_a : x \mapsto \frac{h(x) - h(a)}{x - a}$ est croissante sur $\mathbb{R} \setminus \{a\}$.

Partie I - Préliminaires

On établit ici des résultats qui seront utilisées dans la suite du problème. Les questions sont indépendantes.

1. Soit $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables aléatoires, indépendantes et suivant toutes la loi de Poisson de paramètre $\lambda > 0$. On pose, pour tout entier naturel, non nul $S_n = \sum_{k=1}^n X_k$.

- (a) Quelle est, pour tout entier naturel n non nul, la loi de S_n ?
 (b) Prouver, pour tout $\varepsilon > 0$ et pour tout entier naturel n non nul, l'inégalité :

$$\mathbf{P}\left(\frac{S_n}{n} \geq \lambda + \varepsilon\right) \leq \frac{\lambda}{n\varepsilon^2}.$$

2. Soit $(x, \mu) \in \mathbb{R}^2$ avec $0 < \mu < x$.

On considère l'application Ψ qui, à chaque réel $\theta \geq 0$, associe

$$\Psi(\theta) = \exp\left(\mu(e^\theta - 1) - \theta x\right).$$

- (a) Montrer que la fonction Ψ a un minimum sur $\mathbb{R}_+$ et déterminer sa valeur.
 (b) Soit $\lambda > 0$, $\varepsilon > 0$ et $n \in \mathbb{N}^*$. On pose $x = n(\lambda + \varepsilon)$ et $\mu = n\lambda$.
 Justifier, dans ces conditions, l'existence d'un réel $a > 0$ (fonction de λ et ε mais pas de n) pour lequel $\min_{\theta \geq 0} \Psi(\theta) = e^{-an}$.

3. Soit g une fonction de $\mathbb{R}$ dans $\mathbb{R}$ telle que, pour tout $(x, y) \in \mathbb{R}^2$, $g\left(\frac{x+y}{2}\right) \leq \frac{g(x) + g(y)}{2}$.

- (a) Prouver, pour tout couple $(x, y) \in \mathbb{R}^2$, pour tout entier naturel n et pour tout entier $k \in \llbracket 0, 2^n \rrbracket$, l'inégalité :

$$g\left(\frac{k}{2^n}x + \left(1 - \frac{k}{2^n}\right)y\right) \leq \frac{k}{2^n}g(x) + \left(1 - \frac{k}{2^n}\right)g(y).$$

On pourra raisonner par récurrence sur l'entier n , et observer que, si p est un entier, on a l'égalité : $\frac{2p + (2p+2)}{2} = 2p + 1$.

- (b) On pose $D = \bigcup_{n=0}^{+\infty} \left\{ \frac{k}{2^n} ; k \in \llbracket 0, 2^n \rrbracket \right\}$ et on considère un réel λ de $[0, 1]$.

On note, pour tout entier naturel n , $d_n = \frac{\lfloor \lambda 2^n \rfloor}{2^n}$.

Vérifier que la suite $(d_n)_{n \in \mathbb{N}}$ est une suite d'éléments de D , **croissante** et convergente vers λ .

- (c) On suppose, de plus, que la fonction g est croissante. Montrer qu'elle est convexe.

4. Soit X une variable aléatoire discrète possédant une espérance.

- (a) Rappeler l'inégalité de Markov.

- (b) On suppose, de plus, la variable X à valeurs positives ou nulles. Établir, pour tout réel $x > 0$, l'inégalité

$$\mathbf{P}(X \geq x) \leq \inf_{\theta > 0} \frac{\mathbf{E}(e^{\theta X})}{e^{\theta x}}.$$

5. Soit X une variable aléatoire, **possédant une espérance** et telle que $X(\Omega) = \{x_k; k \in \mathbb{N}^*\}$ où la suite $(x_k)_{k \in \mathbb{N}^*}$ est à valeurs dans $\mathbb{R}_+$, strictement croissante et de limite $+\infty$. On pose $x_0 = 0$. On note, pour tout entier naturel n , $\alpha_n = \mathbf{E}(X \mathbf{1}_{[X \leq x_n]})$.

- (a) Prouver la convergence de la suite $(\alpha_n)_{n \in \mathbb{N}}$.

- (b) Soit $\varepsilon > 0$.

Justifier l'existence d'un entier naturel N tel que, pour tout entier $n \geq N$, on a :

$$\mathbf{E}(X) - \varepsilon \leq \alpha_n.$$

En déduire que la suite $(\alpha_n)_{n \in \mathbb{N}}$ converge vers $\mathbf{E}(X)$.

- (c) Soit β la fonction qui, à chaque réel $K \geq 0$ associe, $\beta(K) = \mathbf{E}(X \mathbf{1}_{[X \leq K]})$.

Prouver l'égalité $\lim_{K \rightarrow +\infty} \beta(K) = \mathbf{E}(X)$.

On admet que ce résultat est valable pour toute variable aléatoire discrète à valeurs positives ou nulles.

6. Soit $f: \mathbb{R} \rightarrow \mathbb{R}$ concave et décroissante. Soit $x \in \mathbb{R}$.

- (a) Justifier l'existence d'une limite finie, négative ou nulle, à droite en x , pour la fonction $u \mapsto \frac{f(u) - f(x)}{u - x}$. On note θ_0 cette limite.

- (b) Prouver, pour tout réel u , l'inégalité : $f(u) \leq f(x) + \theta_0(u - x)$.

- (c) En déduire l'égalité

$$\inf_{\theta \geq 0} \sup_{u \in \mathbb{R}} (f(u) + \theta(u - x)) = f(x),$$

en ayant, pour tout réel θ , posé $\sup_{u \in \mathbb{R}} (f(u) + \theta(u - x)) = +\infty$ quand la fonction $u \mapsto f(u) + \theta(u - x)$ n'est pas majorée.

7. Soit X une variable aléatoire telle que $X(\Omega) = \{x_k; k \in \mathbb{N}^*\}$ où la suite $(x_k)_{k \in \mathbb{N}^*}$ est à valeurs dans $\mathbb{R}_+$, strictement croissante et de limite $+\infty$. On pose $x_0 = 0$.

- (a) Vérifier que la fonction qui à chaque réel t associe $\mathbf{P}(X > t)$ est continue par morceaux sur $\mathbb{R}$.

- (b) Soit r un entier naturel.

Prouver l'égalité $\int_0^{x_{r+1}} \mathbf{P}(X > t) dt = \sum_{k=0}^r (x_{k+1} - x_k) \mathbf{P}(X > x_k)$, puis l'égalité

$$\int_0^{x_{r+1}} \mathbf{P}(X > t) dt = \sum_{k=1}^r x_k \mathbf{P}(X = x_k) + x_{r+1} \mathbf{P}(X > x_r).$$

- (c) Prouver que la variable aléatoire X a une espérance si et seulement si l'intégrale impropre $\int_0^{+\infty} \mathbf{P}(X > t) dt$ converge et que, dans ce cas, on a l'égalité :

$$\mathbf{E}(X) = \int_0^{+\infty} \mathbf{P}(X > t) dt.$$

On admet que ce résultat s'étend à toute variable aléatoire discrète à valeurs positives ou nulles.

8. Soit $(v_n)_{n \in \mathbb{N}^*}$ une suite d'éléments de $\mathbb{R}_+$ vérifiant, pour tout $(m, n) \in (\mathbb{N}^*)^2$, l'inégalité :

$$v_{m+n} \leq v_m + v_n.$$

- (a) Justifier l'existence du nombre $\rho = \inf \left\{ \frac{v_n}{n} ; n \in \mathbb{N}^* \right\}$.

- (b) Soit $\varepsilon > 0$ et $N \in \mathbb{N}^*$ tel que $\rho \leq \frac{v_N}{N} < \rho + \varepsilon$.

Soit $(k, r) \in \mathbb{N} \times \llbracket 0, N-1 \rrbracket$ et $n = kN + r$. Prouver l'inégalité $\frac{v_n}{n} \leq \frac{v_N}{N} + \frac{v_r}{n}$ et en déduire, pour tout entier n assez grand, l'encadrement $\rho \leq \frac{v_n}{n} \leq \rho + 2\varepsilon$.

On en déduit ainsi que $\lim_{n \rightarrow +\infty} \frac{v_n}{n} = \rho$.

Dans toute la suite du problème, on considère une suite $(X_k)_{k \in \mathbb{N}^*}$ de variables aléatoires discrètes, à valeurs **positives ou nulles**, indépendantes et suivant toutes la loi de X_1 . On pose, pour tout entier naturel n non nul, $S_n = \sum_{k=1}^n X_k$.

On suppose, de plus, que, pour tout réel t , $\mathbf{P}(X_1 > t) > 0$.

9. (a) Soit $(m, n) \in (\mathbb{N}^*)^2$ et $(u, v) \in \mathbb{R}^2$. Établir l'inégalité :

$$\mathbf{P}(S_{m+n} \geq mu + nv) \geq \mathbf{P}(S_m \geq mu) \mathbf{P}(S_n \geq nv).$$

- (b) Prouver, pour tout entier naturel n non nul et pour tout réel u , l'inégalité :

$$\mathbf{P}(S_n \geq nu) > 0.$$

- (c) En déduire que, pour tout réel u , $\lim_{n \rightarrow +\infty} \frac{\ln \mathbf{P}(S_n \geq nu)}{n} = \sup_{n \in \mathbb{N}^*} \frac{\ln \mathbf{P}(S_n \geq nu)}{n}$.

Partie II - Le théorème de Cramer

1. Une inégalité

- (a) Soit J une partie de $\mathbb{R}$ telle que, pour tout élément x de J , l'intervalle $]-\infty, x]$ est inclus dans J . Prouver que la partie J est un intervalle de $\mathbb{R}$.
- (b) Vérifier que la fonction $\varphi : \theta \mapsto \ln \mathbf{E}(e^{\theta X_1})$ est définie sur un intervalle I de $\mathbb{R}$ contenant $\mathbb{R}_-$.
- (c) Déterminer l'intervalle I et la valeur de $\varphi(\theta)$ pour tout $\theta \in I$, dans les cas où X_1 suit la loi géométrique de paramètre $p \in]0, 1[$ ou la loi de Poisson de paramètre $\lambda > 0$.

2. (a) Vérifier que la fonction $H : u \mapsto \sup_{n \in \mathbb{N}^*} \frac{1}{n} \ln (\mathbf{P}(S_n \geq nu))$ est définie sur $\mathbb{R}$, et qu'elle est décroissante.

- (b) Prouver, pour tout couple $(u, v) \in \mathbb{R}^2$, l'inégalité : $H\left(\frac{u+v}{2}\right) \geq \frac{1}{2}(H(u) + H(v))$.

On utilisera I - 9.

(c) Qu'en déduit-on pour la fonction H ?

3. Soit $\theta \in I \cap \mathbb{R}_+$.

(a) Justifier, pour tout entier naturel n non nul, l'égalité : $\mathbf{E}(e^{\theta S_n}) = \left(\mathbf{E}(e^{\theta X_1})\right)^n$.

(b) En déduire, pour tout réel u et pour tout entier naturel n non nul, l'inégalité :

$$\varphi(\theta) \geq \frac{1}{n} \ln \left(e^{n\theta u} \mathbf{P}(S_n \geq nu) \right).$$

(c) Conclure à l'inégalité :

$$\varphi(\theta) \geq \sup_{u \in \mathbb{R}} (\theta u + H(u)).$$

4. Le cas d'égalité si $\theta = 0$

Établir l'égalité : $\varphi(0) = \sup_{u \in \mathbb{R}} H(u)$.

5. L'autre inégalité si $\theta > 0$

On rappelle que les variables X_k sont discrètes, à valeurs positives ou nulles, indépendantes et suivent toutes la loi de X_1 , et que, pour tout réel t , $\mathbf{P}(X_1 > t) > 0$.

Soit $K > 0$, $\theta > 0$ et $n \in \mathbb{N}^*$.

(a) i. Prouver l'inégalité :

$$\ln \left(\mathbf{E}(e^{\theta X_1} \mathbf{1}_{[X_1 \leq K]}) \right) = \frac{1}{n} \ln \left(\mathbf{E}(e^{\theta S_n} \prod_{i=1}^n \mathbf{1}_{[X_i \leq K]}) \right).$$

ii. En déduire l'inégalité :

$$\ln \left(\mathbf{E}(e^{\theta X_1} \mathbf{1}_{[X_1 \leq K]}) \right) \leq \frac{1}{n} \ln \left(\mathbf{E}(e^{\theta S_n} \mathbf{1}_{[S_n \leq nK]}) \right).$$

(b) Justifier l'existence d'une espérance pour la variable $e^{\theta S_n} \mathbf{1}_{[S_n \leq nK]}$ et établir l'inégalité :

$$\mathbf{E}(e^{\theta S_n} \mathbf{1}_{[S_n \leq nK]}) \leq \int_0^{\exp(n\theta K)} \mathbf{P}(e^{\theta S_n} > t) dt.$$

(c) En déduire l'inégalité $\mathbf{E}(e^{\theta S_n} \mathbf{1}_{[S_n \leq nK]}) \leq 1 + n\theta \int_0^K \mathbf{P}(S_n > nu) e^{n\theta u} du$, puis l'inégalité :

$$\ln \left(\mathbf{E}(e^{\theta X_1} \mathbf{1}_{[X_1 \leq K]}) \right) \leq \frac{1}{n} \ln \left(1 + n\theta K \exp(nM(\theta)) \right),$$

où l'on a posé $M(\theta) = \sup_{u \in \mathbb{R}} (\theta u + H(u))$.

6. Prouver, pour tout réel K assez grand et pour tout entier naturel n non nul, l'inégalité :

$$\ln \left(\mathbf{E}(e^{\theta X_1} \mathbf{1}_{[X_1 \leq K]}) \right) \leq \frac{\ln(2n\theta K)}{n} + M(\theta).$$

7. En déduire l'inégalité : $\varphi(\theta) \leq M(\theta)$.

Il y a donc égalité via 3 - (c).

8. Conclure, pour tout réel x , à l'égalité :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \mathbf{P}(S_n \geq nx) = \inf_{\theta \geq 0} \left(\ln \mathbf{E}(e^{\theta X_1}) - \theta x \right).$$

9. Soit $\lambda > 0$ et $\varepsilon > 0$.

(a) Justifier l'existence d'un réel $a > 0$ pour lequel, pour tout entier naturel n assez grand, on a :

$$\mathbf{P}\left(\frac{S_n}{n} \geq \lambda + \varepsilon\right) \leq e^{-an}.$$

(b) Retrouver ce résultat, en utilisant les résultats des questions I - 2 et I - 4, dans le cas particulier où X_1 suit la loi de Poisson de paramètre $\lambda > 0$.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2019

Épreuve à option (A) : Mathématiques

Durée : 4h

L'objet du problème est l'étude de quelques fonctions périodiques, bâties à partir de la distance à une partie de $\mathbb{R}$ et utilisées notamment en théorie du signal et en théorie des probabilités.

- Pour toute partie A de $\mathbb{R}$, on note $-A$ l'ensemble des opposés des éléments de A :

$$-A = \{-a; a \in A\} .$$

- Soit f une application de $\mathbb{R}$ dans $\mathbb{R}$.

Pour tout réel strictement positif t , on dit que la fonction f est *t-périodique* si elle vérifie :

$$\forall x \in \mathbb{R}, f(x+t) = f(x) .$$

On dit que f est périodique s'il existe un réel strictement positif t pour lequel la fonction f est t -périodique. Dans ce cas, on appelle *période* de f la borne inférieure de l'ensemble des réels strictement positifs t pour lesquels la fonction f est t -périodique.

Par exemple, la fonction $x \longmapsto x - \lfloor x \rfloor$ est périodique, de période égale à 1, et toute fonction constante est périodique, de période égale à 0.

Les quatre parties du problème sont largement indépendantes, mais les matrices introduites dans la partie I sont utilisées dans la partie IV.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Partie I : une famille de matrices de $\mathcal{M}_3(\mathbb{R})$

Pour tout $c \in \mathbb{R}$, on note T_c la matrice de $\mathcal{M}_3(\mathbb{R})$ donnée par :

$$T_c = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & c \end{pmatrix}.$$

On considère d'autre part la matrice : $S = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

1. a) Pour quelles valeurs de c la matrice T_c est-elle diagonalisable?
 b) Existe-t-il une valeur de c et une matrice inversible P pour lesquelles les deux matrices $P^{-1}T_cP$ et $P^{-1}SP$ sont diagonales?

2. Soit $n \in \mathbb{N}^*$.

a) Calculer la matrice $(T_{1/2})^n$.

b) En déduire, pour tout entier $m \in \mathbb{N}$, la matrice $S^m (T_{1/2})^n$.

c) Pour tout $c \in \mathbb{R}$, on note $\mathcal{G}$ le plus petit sous-ensemble de $\mathcal{M}_3(\mathbb{R})$ stable par le produit matriciel et contenant les matrices S et $T_{1/2}$.

Justifier que toutes les matrices de $\mathcal{G}$ sont de la forme $2^{-n} \begin{pmatrix} 2^n & h & k \\ 0 & (-1)^m & \ell \\ 0 & 0 & 1 \end{pmatrix}$ avec $(h, k, \ell) \in \mathbb{Z}^3$

et $(m, n) \in \mathbb{N}^2$.

d) Quelles sont les matrices de $\mathcal{G}$ dont les coefficients diagonaux sont égaux à $+1$ ou à -1 ?

Partie II : distances à une partie de $\mathbb{R}$

3. a) Justifier, pour tout réel x et pour toute partie non vide A de $\mathbb{R}$, l'existence de la distance de x à A , c'est-à-dire du nombre réel positif ou nul

$$d(x, A) = \inf\{|x - a|; a \in A\}.$$

b) Quels sont les nombres réels x pour lesquels la distance $d(x, A)$ de x à une partie non vide A de $\mathbb{R}$ est nulle?

c) Justifier que, si A est une partie fermée de $\mathbb{R}$, alors il existe pour tout réel x un élément a_x de A tel que :

$$|x - a_x| = d(x, A).$$

4. Pour toute partie non vide A de $\mathbb{R}$, on note d_A la fonction $x \mapsto d(x, A)$, définie sur $\mathbb{R}$.

a) Dessiner le graphe de la fonction d_A lorsque $A = \{-1, 0, +1\}$.

b) Démontrer que si $A = -A$, alors la fonction d_A est paire et que la réciproque est vraie lorsque A est fermée.

c) Démontrer que, si l'ensemble A est minoré ou majoré, alors la fonction d_A n'est pas bornée.

d) Donner un exemple d'ensemble A qui n'est ni minoré ni majoré, pour lequel la fonction d_A n'est pas bornée.

5. Soit A une partie non vide de $\mathbb{R}$.

a) Démontrer que, pour tout couple $(x, y) \in \mathbb{R}^2$ et tout $\varepsilon > 0$, on a :

$$d_A(y) \leq d_A(x) + |y - x| + \varepsilon.$$

b) En déduire que la fonction d_A est continue sur $\mathbb{R}$.

c) Justifier que, si la fonction d_A est dérivable et nulle en un point x_0 , alors sa dérivée en ce point est nécessairement nulle.

6. On considère dans cette question une suite strictement décroissante $(u_n)_{n \in \mathbb{N}^*}$, convergente et de limite nulle, et on note :

$$A = \mathbb{R}_-^* \cup \{u_n; n \in \mathbb{N}^*\}.$$

a) Justifier que 0 appartient à la frontière de A et que la fonction d_A s'annule en 0.

b) On considère ici le cas où $u_n = \frac{1}{n!}$ pour tout $n \in \mathbb{N}^*$.

Calculer $d_A\left(\frac{1}{2}(u_n + u_{n+1})\right)$ et en déduire que la fonction d_A n'est pas dérivable à droite au point 0.

c) On considère maintenant le cas où $u_n = \frac{1}{n}$ pour tout $n \in \mathbb{N}^*$.

Démontrer que la fonction d_A est dérivable au point 0 et indiquer les autres points où elle est dérivable.

Partie III : signaux triangulaires

7. Dans cette question, on suppose que A est un sous-groupe du groupe additif $(\mathbb{R}, +)$.

a) Donner la fonction d_A lorsque $A = \{0\}$ et lorsque A est dense dans $\mathbb{R}$.

b) Démontrer que, dans tous les autres cas, il existe un réel strictement positif α tel que :

$$\forall x \in \mathbb{R}, \quad d_A(x) = \alpha d_{\mathbb{Z}}\left(\frac{x}{\alpha}\right).$$

On note désormais σ la fonction $d_{\mathbb{Z}}$:

$$\forall x \in \mathbb{R}, \quad \sigma(x) = \min\{|x - n|; n \in \mathbb{Z}\} \quad (1)$$

8. a) Justifier l'expression de $d_{\mathbb{Z}}(x)$ donnée par (1).

b) Vérifier que la fonction σ est paire, périodique et de période égale à 1.

c) Calculer l'intégrale $\int_{-1/2}^{+1/2} \sigma(t) dt$.

d) Démontrer que, pour tout entier $n \in \mathbb{N}^*$, on a :

$$\int_{-1/2}^{+1/2} \sigma(t) \cos(2\pi n t) dt = \begin{cases} -\frac{1}{n^2 \pi^2} & \text{si } n \text{ est impair} \\ 0 & \text{si } n \text{ est pair} \end{cases}$$

Pour tout $r \in]0, 1[$, on note P_r l'application $t \mapsto \frac{1 - r^2}{1 - 2r \cos(2\pi t) + r^2}$.

9. a) Vérifier que, pour tout $r \in [0, 1[$, la fonction P_r est définie, périodique et continue sur $\mathbb{R}$.

b) Soit $\varepsilon \in]0, 1/2[$. Démontrer que l'intégrale $\int_{\varepsilon}^{+1/2} P_r(t) dt$ tend vers 0 quand r tend vers 1 par valeurs inférieures.

10. Soit $t \in \mathbb{R}$.

a) Démontrer que le rayon de convergence de la série entière $\sum_{n \geq 1} (\cos(2\pi n t)) x^n$ est égal à 1.

b) Établir, pour tout $r \in [0, 1[$, l'égalité

$$P_r(t) = 1 + 2 \sum_{n=1}^{+\infty} (\cos(2\pi n t)) r^n.$$

11. Soit $x \in \mathbb{R}$.

Pour tout $r \in [0, 1]$, on pose : $\sigma_r(x) = \frac{1}{4} - \frac{2}{\pi^2} \sum_{k=0}^{+\infty} \frac{\cos((2k+1)2\pi x)}{(2k+1)^2} r^{2k+1}$.

a) En utilisant les formules démontrées en 8c et 8d, démontrer que :

$$\forall r \in [0, 1[, \quad \sigma_r(x) = \int_{-1/2}^{+1/2} \sigma(t) \left(1 + 2 \sum_{n=1}^{+\infty} \cos(2\pi n x) \cos(2\pi n t) r^n\right) dt.$$

b) En déduire, pour tout $r \in [0, 1[$, les égalités :

$$\sigma_r(x) = \frac{1}{2} \left(\int_{-1/2}^{+1/2} \sigma(t) P_r(x+t) dt + \int_{-1/2}^{+1/2} \sigma(t) P_r(x-t) dt \right) = \int_{-1/2}^{+1/2} \sigma(x+t) P_r(t) dt.$$

12. Déduire de ce qui précède que pour tout $x \in \mathbb{R}$, on a :

$$\sigma(x) = \frac{1}{4} - \frac{2}{\pi^2} \sum_{n=0}^{+\infty} \frac{\cos((2n+1)2\pi x)}{(2n+1)^2}.$$

Partie IV : une famille de fonctions 1-périodiques

Pour tout réel $c \in]0, 1[$, on note τ_c l'application

$$\tau_c : x \longmapsto \sum_{n=0}^{+\infty} c^n \sigma(2^n x) \quad (2)$$

où σ est la fonction définie par (1) dans la partie III.

13. Soit $c \in]0, 1[$.

- a) Justifier que la fonction τ_c est définie, continue et 1-périodique sur $\mathbb{R}$.
- b) Vérifier que, pour tout $x \in \mathbb{R}$: $\tau_c(x) = \sigma(x) + c\tau_c(2x)$.
- c) Démontrer que τ_c est l'unique fonction τ définie et bornée sur $\mathbb{R}$ vérifiant :

$$\forall x \in \mathbb{R}, \quad \tau(x) = \sigma(x) + c\tau(2x) \quad (3)$$

14. Dans cette question, on suppose que c est égal à $1/4$.

- a) Démontrer que, pour tout $x \in [0, 1/2]$, on a : $\tau_c(x) = 2x(1-x)$.
- b) Donner une représentation graphique de la fonction τ_c .

On note désormais β la fonction $\tau_{1/2}$:

$$\forall x \in \mathbb{R}, \quad \beta(x) = \sum_{n=0}^{+\infty} 2^{-n} \sigma(2^n x) \quad (4)$$

15. Soit E le $\mathbb{R}$ -espace vectoriel des applications continues de $[0, 1]$ dans $\mathbb{R}$.

On note f_1, f_2, f_3 les restrictions à $[0, 1]$ des trois fonctions $x \mapsto 1$, $x \mapsto x$, $x \mapsto \beta(x)$, respectivement, et F le sous-espace vectoriel de E engendré par ces trois fonctions.

On note Φ et Ψ les deux applications définies sur E par :

$$\forall f \in E, \quad \forall x \in [0, 1], \quad \begin{cases} \Phi(f)(x) = f(x/2) \\ \Psi(f)(x) = f(1-x) \end{cases}.$$

- a) Vérifier que le sous-espace vectoriel F est stable par les endomorphismes Φ et Ψ .
- b) Vérifier que (f_1, f_2, f_3) est une base de F et reconnaître les matrices dans cette base des endomorphismes de F induits par Φ et Ψ .
- c) Soit $n \in \mathbb{N}$ et $x \in [0, 2^{-n}]$.

En utilisant le résultat de la question 2b de la partie I, exprimer $\beta((1-x)2^{-n})$ à l'aide de x , n et $\beta(x)$.

- 16. a) En exploitant le résultat trouvé en 15 c, justifier que la fonction β n'est pas dérivable en 0.
- b) Démontrer qu'il n'existe aucun intervalle ouvert non vide de $\mathbb{R}$ sur lequel la fonction β est dérivable.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2019

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet se compose d'un exercice et d'un problème dont les différentes parties peuvent, dans une certaine mesure, être traitées de manière indépendante, quitte à admettre les résultats des parties précédentes. La qualité de la rédaction et la précision des réponses apportées entrent pour une part importante dans la notation. Chaque réponse doit être suffisamment justifiée. Les réponses absurdes sont pénalisées.

Problème

On considère un entier $n \geq 2$, une succession de $n + 1$ dates $t_0 < t_1 < \dots < t_n$, et un titre financier dont les valeurs successives (en €) lors des dates $t_0, \dots, t_n$ sont respectivement notées $S_0, \dots, S_n$.

On suppose que, sur chaque intervalle de temps $[t_i, t_{i+1}[$ (où $0 \leq i \leq n - 1$), un individu détient à la fois une certaine somme c_i (en €) déposée sur un compte bancaire, et une certaine quantité d_i du titre financier. A chaque date intermédiaire t_i , l'individu a la possibilité de modifier la répartition de ses avoirs entre compte bancaire et titre financier, et l'on suppose qu'il n'y a ni retrait ni apport de fonds lors d'une telle modification, si bien que l'on a, pour tout $0 \leq i \leq n - 2$, la relation :

$$c_i + d_i S_{i+1} = c_{i+1} + d_{i+1} S_{i+1}. \quad (1)$$

On suppose en outre que toutes les valeurs $c_i \in \mathbb{R}$ et $d_i \in \mathbb{R}$ sont possibles (y compris des valeurs négatives pour c_i et d_i , qui correspondent concrètement au cas où l'individu a emprunté une somme en € ou une certaine quantité de titres). On notera R_i l'avoir total de l'individu (en €) à la date t_i , soit $R_0 = c_0 + d_0 S_0$ et, pour tout $1 \leq i \leq n$,

$$R_i = c_{i-1} + d_{i-1} S_i. \quad (2)$$

Pour décrire l'évolution de la valeur du titre au cours du temps, on introduit une suite de variables aléatoires $W_1, \dots, W_n$, mutuellement indépendantes, et possédant toutes la même loi de probabilité donnée par :

$$\mathbb{P}(W_i = h) = p, \quad \mathbb{P}(W_i = b) = 1 - p,$$

h, b et p étant des nombres réels fixés vérifiant les conditions $0 < b < 1 < h$ et $0 < p < 1$. On suppose que $S_0 \equiv 1$ €, et l'on modélise l'évolution de la valeur du titre au cours du temps en supposant que celle-ci obéit, pour tout $1 \leq i \leq n$, à la relation :

$$S_i = W_i \cdot S_{i-1}.$$

Partie I – Quelques calculs préliminaires

- 1) Exprimer S_n en fonction des variables aléatoires $W_1, \dots, W_n$.
- 2) Exprimer la valeur de l'espérance $\mathbb{E}(S_n)$ en fonction de n, p, b, h .
- 3) Exprimer la valeur de la variance $\mathbb{V}(S_n)$ en fonction de n, p, b, h .
- 4) On considère une fonction $f :]0, +\infty[\rightarrow \mathbb{R}$. Montrer qu'il existe un unique couple de fonctions $c_f :]0, +\infty[\rightarrow \mathbb{R}$ et $d_f :]0, +\infty[\rightarrow \mathbb{R}$, telles que, pour tout $s \in]0, +\infty[$, le système d'équations suivant est satisfait :

$$\begin{cases} c_f(s) + s \cdot h \cdot d_f(s) = f(h \cdot s) \\ c_f(s) + s \cdot b \cdot d_f(s) = f(b \cdot s) \end{cases}$$

Une expression explicite des fonctions c_f et d_f à l'aide de f, h et b est demandée.

- 5) Montrer que, avec les notations de la question 4), on a, pour tout s :

$$c_f(s) + s \cdot d_f(s) = \left(\frac{1-b}{h-b} \right) \cdot f(hs) + \left(\frac{h-1}{h-b} \right) \cdot f(bs).$$

Partie II – Réplication avec horizon fixe

Dans toute cette partie, on considère une fonction $g :]0, +\infty[\rightarrow \mathbb{R}$ fixée.

6) On suppose que, à la date t_{n-1} , on dispose de $c_{n-1} = c_g(S_{n-1}) \in$ sur le compte bancaire, et d'une quantité $d_{n-1} = d_g(S_{n-1})$ du titre financier (les fonctions c_g et d_g sont celles définies à la question 4) pour $f = g$). Montrer que, à la date t_n , la richesse totale de l'individu vérifie nécessairement $R_n = g(S_n)$. (On prêtera bien attention au fait que, ici, S_{n-1} et S_n sont des variables aléatoires.)

7) On définit récursivement les fonctions $f_0, \dots, f_n :]0, +\infty[\rightarrow \mathbb{R}$ de la manière suivante : on pose d'abord $f_n = g$, puis, pour $0 \leq i \leq n-1$,

$$f_i(s) = c_{f_{i+1}}(s) + s d_{f_{i+1}}(s). \quad (3)$$

On suppose que, à la date t_0 , l'individu dispose d'un avoir initial de $f_0(S_0) \in$, soit $R_0 = f_0(S_0)$. Montrer que, en répartissant son avoir de manière appropriée à chaque nouvelle date, il lui est possible de maintenir la validité de la relation $R_i = f_i(S_i)$ pour tout $1 \leq i \leq n$.

Indication : Pour commencer, l'individu répartit son avoir initial en $c_0 = c_{f_1}(S_0) \in$ sur le compte bancaire et une quantité $d_0 = d_{f_1}(S_0)$ du titre financier.

8) **Application :** Alice souhaiterait (pour des raisons qui lui sont propres) pouvoir disposer d'une somme de $g(S_n) \in$ à la date t_n , en échange d'un versement initial à la date t_0 , sans avoir à s'occuper d'opérations intermédiaires de gestion entre ces deux dates. Son ami Barnabé s'engage alors à lui fournir la somme de $g(S_n) \in$ à la date t_n en échange du versement de $v_0 \in$ à la date t_0 . En utilisant les résultats des questions précédentes, expliquer comment, en fixant $v_0 = f_0(S_0)$, Barnabé parviendra à fournir à Alice la somme de $g(S_n) \in$ à la date t_n , sans gain ni perte pour lui.

On introduit maintenant une probabilité $\mathbb{Q}$ distincte (en général) de la probabilité de référence $\mathbb{P}$ considérée jusqu'ici, et l'on suppose que, sous la probabilité $\mathbb{Q}$, les variables aléatoires $W_1, \dots, W_n$ sont mutuellement indépendantes et possèdent toutes la même loi de probabilité donnée par :

$$\mathbb{Q}(W_i = h) = \frac{1-b}{h-b}, \quad \mathbb{Q}(W_i = b) = \frac{h-1}{h-b}.$$

L'espérance calculée par rapport à la probabilité $\mathbb{Q}$ sera notée $\mathbb{E}_{\mathbb{Q}}$ dans la suite, pour la distinguer de l'espérance $\mathbb{E}$ calculée par rapport à la probabilité $\mathbb{P}$.

9) Vérifier que les valeurs de $\mathbb{Q}(W_i = h)$ et $\mathbb{Q}(W_i = b)$ données ci-dessus constituent une spécification correcte de la loi de probabilité d'une variable aléatoire ne prenant que les deux valeurs h et b .

10) Pour tout $0 \leq i \leq n-1$, on définit la fonction $\phi_i :]0, +\infty[\rightarrow \mathbb{R}$ par la formule :

$$\phi_i(s) = \mathbb{E}_{\mathbb{Q}} [g(s \cdot W_{i+1} \cdots W_n)], \quad (4)$$

et l'on pose en outre $\phi_n = g$. Montrer que, pour tout $1 \leq i \leq n-1$, on a la relation :

$$\phi_i(s) = \left(\frac{1-b}{h-b} \right) \phi_{i+1}(sh) + \left(\frac{h-1}{h-b} \right) \phi_{i+1}(sb).$$

Indication : utiliser une décomposition selon les différentes valeurs possibles de W_{i+1} .

11) Dédire de ce qui précède que, pour tout $0 \leq i \leq n$, on a $\phi_i = f_i$.

Indication : relire la question 5).

- 12) En déduire la relation (importante!) $f_0(S_0) = \mathbb{E}_{\mathbb{Q}}(g(S_n))$.
- 13) On introduit la variable aléatoire B définie comme le nombre d'indices $i \in \{1, \dots, n\}$ tels que $W_i = h$.
- (a) Quelle est la loi de B (par rapport à la probabilité $\mathbb{Q}$) ?
 - (b) Exprimer S_n comme une fonction de B et de h, b, n .
 - (c) En déduire une formule explicite (sous forme d'une somme) pour $f_0(S_0)$ en fonction de h, b, n et g .
- 14) Montrer l'identité :

$$f_0(S_0) = \mathbb{E} \left[g(S_n) \cdot \left(\frac{1}{p} \right)^B \left(\frac{1-b}{h-b} \right)^B \left(\frac{1}{1-p} \right)^{n-B} \left(\frac{h-1}{h-b} \right)^{n-B} \right].$$

- 15) Commenter le fait que les raisonnements précédents conduisent à choisir $v_0 = \mathbb{E}_{\mathbb{Q}}(g(S_n))$ et non pas $v_0 = \mathbb{E}(g(S_n))$, ou encore une valeur de v_0 caractérisée (par exemple) par une condition du type $\mathbb{P}(g(S_n) > v_0) \leq \epsilon$ pour une « petite » valeur de ϵ .

Partie III – Réplication avec horizon variable

Alice souhaite conclure un accord différent de celui étudié à la question 8), qui lui permettrait de choisir à tout moment la date à laquelle elle recevrait le versement de Barnabé. En échange d'un montant initial de $z_0 \in$ à la date t_0 , Alice pourrait donc décider à la date t_k de demander à Barnabé de lui verser la somme de $g(S_k) \in$ (Alice ne pouvant exiger qu'une seule fois le versement, parmi l'ensemble des dates possibles.)

- 16) A priori, la valeur de z_0 devrait-elle être supérieure, égale, ou inférieure à la valeur v_0 considérée à la question 8) ? (Justifier la réponse.)

On définit récursivement les fonctions $\psi_0, \dots, \psi_n :]0, +\infty[\rightarrow \mathbb{R}$ de la manière suivante : on pose d'abord $\psi_n = g$, puis, pour $0 \leq i \leq n-1$,

$$\psi_i(s) = \max(g(s), c_{\psi_{i+1}}(s) + sd_{\psi_{i+1}}(s)). \quad (5)$$

- 17) Expliquer comment, en fixant $z_0 = \psi_0(S_0)$, Barnabé parviendra à disposer à toute date t_k d'une somme au moins égale à $g(S_k) \in$ jusqu'à ce qu'Alice en fasse la demande.
- 18) Expliquer pourquoi, si Alice demande à Barnabé le versement de $g(S_k) \in$ à une date t_k pour laquelle on a $g(S_k) < \psi_k(S_k)$, Barnabé dégagera un profit de cette opération. Pour Alice, quelle pourrait-être la meilleure date à laquelle demander le versement ?

Partie IV – Un petit contre-exemple

Dans cette partie, on considère simplement deux dates t_0 et t_1 , et l'on suppose que W_1 peut prendre trois valeurs distinctes : $\mathbb{P}(W_1 = 0,5) = 0,2$; $\mathbb{P}(W_1 = 1) = 0,6$; $\mathbb{P}(W_1 = 2) = 0,2$, tandis que la fonction g vérifie $g(0,5) = 1$; $g(1) = 2$; $g(2) = 3$.

- 19) Vérifier qu'il n'existe pas de couple de nombres réels c, d tels que $\mathbb{P}(c + dS_1 = g(S_1)) = 1$.
- 20) Un accord entre Alice et Barnabé du type de celui considéré à la question 8) est-il possible ici ? (Justifier la réponse.)

Exercice

Dans cet exercice, on considère un nombre entier $N \geq 1$, et N variables aléatoires $U_1, \dots, U_N$, supposées **mutuellement indépendantes, et suivant chacune la loi uniforme sur $]0, 1[$** . On rappelle que la loi uniforme sur l'intervalle $]0, 1[$ est caractérisée par la fonction de densité f_U définie sur $\mathbb{R}$ par $f_U(u) = 1$ pour $u \in]0, 1[$, et $f_U(u) = 0$ pour $u \notin]0, 1[$.

On définit ensuite les variables aléatoires $U_{(1)}, \dots, U_{(N)}$ obtenues en **triant par ordre croissant** les valeurs prises par les variables aléatoires¹ $U_1, \dots, U_N$.

Par exemple, avec $N = 4$, si les valeurs effectivement réalisées des variables aléatoires U_1, U_2, U_3, U_4 se trouvent être :

$$U_1(\omega) = 0,44; U_2(\omega) = 0,72; U_3(\omega) = 0,23; U_4(\omega) = 0,44,$$

on aura :

$$U_{(1)}(\omega) = 0,23; U_{(2)}(\omega) = 0,44; U_{(3)}(\omega) = 0,44; U_{(4)}(\omega) = 0,72.$$

L'objectif de cet exercice est de caractériser la loi de probabilité de la variable $U_{(m)}$, pour chaque $m \in \{1, \dots, N\}$.

A) Calculer la valeur de $\mathbb{E}(U_m)$ pour $1 \leq m \leq N$.

B) Etant donné $x \in [0, 1]$, on appelle $N(x)$ la variable aléatoire définie comme le nombre d'indices $i \in \{1, \dots, N\}$ tels que $U_i \leq x$. Montrer que, pour tout entier $m \in \{1, \dots, N\}$, et tout $x \in [0, 1]$, la réalisation de l'événement $U_{(m)} \leq x$ est équivalente à la réalisation de l'événement $N(x) \geq m$, et en déduire l'identité :

$$\mathbb{P}(U_{(m)} \leq x) = \mathbb{P}(N(x) \geq m).$$

C) Montrer que, pour tout $x \in [0, 1]$, la variable aléatoire $N(x)$ suit la loi binomiale $\mathcal{B}(N, p)$ avec $p = x$, et en déduire que, pour tout $m \in \{1, \dots, N\}$,

$$\mathbb{P}(U_{(m)} \leq x) = \sum_{k=m}^N \binom{N}{k} x^k (1-x)^{N-k}.$$

D) En déduire que, pour tout $m \in \{1, \dots, N\}$, la loi de la variable aléatoire $U_{(m)}$ possède une densité $f_{m,N}$ définie sur $\mathbb{R}$ par :

$$f_{m,N}(x) = \begin{cases} \frac{N!}{(m-1)!(N-m)!} x^{m-1} (1-x)^{N-m} & \text{pour } x \in]0, 1[\\ 0 & \text{pour } x \notin]0, 1[\end{cases}.$$

E) A l'aide d'une intégration par parties, montrer l'identité :

$$\int_0^1 x f_{m,N}(x) dx = \frac{m}{N-m+1} \left(\int_0^1 f_{m,N}(x) dx - \int_0^1 x f_{m,N}(x) dx \right).$$

F) En déduire que, pour tout $m \in \{1, \dots, N\}$:

$$\mathbb{E}(U_{(m)}) = \frac{m}{N+1}.$$

G) Commenter la différence entre les valeurs de $\mathbb{E}(U_m)$ et $\mathbb{E}(U_{(m)})$.

1. De manière plus formelle, pour tout $\omega \in \Omega$, les valeurs $U_{(1)}(\omega), \dots, U_{(N)}(\omega)$ sont obtenues en triant dans l'ordre croissant les valeurs $U_1(\omega), \dots, U_N(\omega)$, une valeur apparaissant plusieurs fois étant alors répétée avec sa multiplicité.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2020

Épreuve de mathématiques

Durée : 3h

En dehors de sa dernière question, la partie II est indépendante de la partie I.

Partie I - Intégrales généralisées de Dirichlet

1. Les sous-questions sont indépendantes.

(a) On considère la fonction

$$g: \mathbb{R} \longrightarrow \mathbb{R} \quad t \longmapsto \begin{cases} \frac{\sin t}{t} & \text{si } t \neq 0; \\ 1 & \text{si } t = 0. \end{cases}$$

Montrer que la fonction g est de classe C^1 sur $\mathbb{R}$.

(b) i. Prouver la convergence de l'intégrale impropre $\int_0^{+\infty} \frac{\sin t}{t} dt$.

On **admet** l'égalité $\int_0^{+\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}$.

ii. Déterminer, pour tout $j \in \mathbb{N}^*$, la valeur de l'intégrale impropre $\int_0^{+\infty} \frac{\sin(jt)}{t} dt$.

(c) Donner le développement limité en 0 à l'ordre 4 de la fonction $t \mapsto \ln g(t)$.

(d) On rappelle que $\int_0^{+\infty} e^{-\frac{u^2}{2}} du = \sqrt{\frac{\pi}{2}}$.

Donner, lorsque l'entier n tend vers $+\infty$, un équivalent de $\int_0^{\frac{\ln n}{\sqrt{n}}} e^{-\frac{nt^2}{6}} dt$.

2. Soit n un entier naturel au moins égal à 2.

(a) Vérifier que l'intégrale impropre $\int_0^{+\infty} \frac{\sin^n t}{t^n} dt$ est convergente.

(b) Quelle est la valeur de $\int_0^{+\infty} \frac{\sin^2 t}{t^2} dt$?

3. On pose, pour tout entier naturel n non nul et tout réel $t > 0$, $h_n(t) = \sin^n t$.

Soit n un entier naturel au moins égal à 2.

(a) Soit $k \in \llbracket 0, n-1 \rrbracket$. Justifier l'existence d'un réel $K > 0$ pour lequel, pour tout réel t , on a : $|h_n^{(k)}(t)| \leq K$.

(b) i. Quel est le développement limité à l'ordre n en 0 de la fonction h_n ?

ii. Établir l'égalité : $\lim_{t \rightarrow 0} \frac{h_n^{(k)}(t)}{t^{n-k}} = \frac{n!}{(n-k)!}$.

(c) Justifier, pour tout entier $k \in \llbracket 0, n-2 \rrbracket$, la convergence absolue de l'intégrale $\int_0^{+\infty} \frac{h_n^{(k)}(t)}{t^{n-k}} dt$.

(d) Justifier la convergence de l'intégrale $\int_0^{+\infty} \frac{h_n^{(n-1)}(t)}{t} dt$ et établir l'égalité

$$\int_0^{+\infty} \left(\frac{\sin t}{t}\right)^n dt = \frac{1}{(n-1)!} \int_0^{+\infty} \frac{h_n^{(n-1)}(t)}{t} dt.$$

4. Soit $n \in \mathbb{N}^*$.

(a) Établir, pour tout réel t , l'égalité

$$h_{2n}(t) = \sin^{2n} t = \frac{1}{4^n} \sum_{k=0}^{2n} (-1)^{n+k} \binom{2n}{k} e^{i(2n-2k)t}.$$

(b) En déduire, pour tout réel t , l'égalité

$$h_{2n}^{(2n-1)}(t) = \sum_{j=1}^n (-1)^{n+j} \binom{2n}{n+j} j^{2n-1} \sin(2jt).$$

(c) En déduire l'égalité : $\int_0^{+\infty} \left(\frac{\sin t}{t}\right)^{2n} dt = \frac{\pi}{2} \frac{1}{(2n-1)!} \sum_{j=1}^n (-1)^{n+j} \binom{2n}{n+j} j^{2n-1}.$

5. Étude asymptotique de la suite de terme général $\int_0^{+\infty} \left(\frac{\sin t}{t}\right)^n dt$

(a) Prouver, quand l'entier n tend vers $+\infty$, l'évaluation $\int_{\frac{\pi}{2}}^{+\infty} \left(\frac{\sin t}{t}\right)^n dt = o\left(\frac{1}{\sqrt{n}}\right).$

(b) i. Étudier la monotonie de la fonction $t \mapsto \frac{\sin t}{t}$ sur l'intervalle $]0, \frac{\pi}{2}]$.

ii. En déduire, quand l'entier n tend vers $+\infty$, l'évaluation asymptotique :

$$\int_{\frac{\ln n}{\sqrt{n}}}^{\frac{\pi}{2}} \left(\frac{\sin t}{t}\right)^n dt = o\left(\frac{1}{\sqrt{n}}\right).$$

On pourra, en utilisant le résultat de la question 1-c), donner un équivalent de $\ln\left(\frac{\sin \varepsilon_n}{\varepsilon_n}\right)$ où $\varepsilon_n = \frac{\ln n}{\sqrt{n}}.$

(c) i. Justifier l'existence de $a > 0$ tel que, pour tout $u \in [0, a]$, on a : $|e^{-u} - 1| \leq 2u$.
ii. Justifier l'existence d'un réel $b > 0$ tel que, pour tout $t \in]0, b]$,

$$-t^3 \leq \ln\left(\frac{\sin t}{t}\right) + \frac{t^2}{6} \leq 0.$$

On utilisera le résultat de la question 1-c).

iii. En déduire, pour tout entier n assez grand, l'inégalité

$$\left| \int_0^{\frac{\ln n}{\sqrt{n}}} \left(\frac{\sin t}{t} \right)^n dt - \int_0^{\frac{\ln n}{\sqrt{n}}} e^{-\frac{nt^2}{6}} dt \right| \leq \int_0^{\frac{\ln n}{\sqrt{n}}} (1 - e^{-nt^3}) dt,$$

puis, toujours quand l'entier n est assez grand, l'inégalité

$$\left| \int_0^{\frac{\ln n}{\sqrt{n}}} \left(\frac{\sin t}{t} \right)^n dt - \int_0^{\frac{\ln n}{\sqrt{n}}} e^{-\frac{nt^2}{6}} dt \right| \leq 2 \frac{\ln^4 n}{n}.$$

(d) En déduire, lorsque l'entier n tend vers $+\infty$, l'équivalence :

$$\int_0^{+\infty} \left(\frac{\sin t}{t} \right)^n dt \sim \sqrt{\frac{3\pi}{2n}}.$$

On se souviendra du résultat de 1-d).

Partie II Montées d'une permutation de $\llbracket 1, n \rrbracket$

On appelle, pour tout entier naturel n non nul, **montée** d'une liste $a = (a_1, a_2, \dots, a_n)$ d'entiers naturels **distincts deux à deux** toute sous-liste $(a_p, a_{p+1}, \dots, a_q)$ (avec $p \leq q$) vérifiant les conditions :

$$\begin{cases} p = 1 \text{ ou } a_{p-1} > a_p \\ \text{et } a_p < a_{p+1} < \dots < a_q \text{ (si } p < q) \\ \text{et } q = n \text{ ou } a_q > a_{q+1} \end{cases}$$

On note $M(a)$ le nombre de montées de la liste a . Par exemple, les montées de la liste $a = (2, 5, 7, 6, 1, 4, 3, 8)$ sont $(2, 5, 7)$, (6) , $(1, 4)$ et $(3, 8)$, et donc $M(a) = 4$.

On définit de même la notion de descente d'une liste a d'entiers naturels distincts deux à deux et son nombre de descentes $D(a)$. Par exemple, les descentes de la liste $a = (2, 5, 7, 6, 1, 4, 3, 8)$ sont (2) , (5) , $(7, 6, 1)$, $(4, 3)$ et (8) , et donc $D(a) = 5$.

On note, pour tout entier naturel n non nul, S_n l'ensemble des n -listes d'éléments de $\llbracket 1, n \rrbracket$ **distincts deux à deux**.

Il est clair que, pour toute liste a de S_n , on a $1 \leq M(a) \leq n$ et $1 \leq D(a) \leq n$.

Enfin, pour tout entier naturel k , on note $E_n(k)$ le nombre de listes a de S_n ayant exactement k montées. Autrement dit, $E_n(k) = \text{card} \{ a \in S_n ; M(a) = k \}$. On a donc, $E_n(0) = 0$ ainsi que $E_n(k) = 0$ pour tout entier $k > n$.

1. Soit $n \in \mathbb{N}^*$.

(a) Déterminer les valeurs de $E_n(1)$ et $E_n(n)$.

(b) Soit k un entier fixé compris entre 1 et n .

Donner un exemple de liste a de S_n pour laquelle $M(a) = k$.

2. Déterminer, pour tout entier naturel n non nul et pour toute liste a de S_n , la valeur de $M(a) + D(a)$.

En notant, pour $i \in \llbracket 1, n-1 \rrbracket$, s_i la somme du nombre de montées et du nombre de descentes de $(a_1, a_2, \dots, a_i)$, on évaluera, en fonction du nombre s_i , la somme s_{i+1} des nombres de montées et de descentes de $(a_1, a_2, \dots, a_{i+1})$.

3. Soit $n \in \mathbb{N}^*$. On associe, à toute liste $a = (a_1, a_2, \dots, a_n)$ de S_n , la liste

$$\Psi(a) = (n+1-a_1, n+1-a_2, \dots, n+1-a_n).$$

- (a) Vérifier que l'application Ψ est une bijection de S_n sur S_n .
- (b) En déduire, pour tout entier k de $\llbracket 1, n \rrbracket$, l'égalité : $E_n(k) = E_n(n+1-k)$.

4. Calcul de $E_n(2)$

Soit n un entier naturel au moins égal à 2.

- (a) Quel est le nombre de couples (A, B) de parties **non vides** de $\llbracket 1, n \rrbracket$ tels que

$$A \cup B = \llbracket 1, n \rrbracket \text{ et } A \cap B = \emptyset.$$

- (b) Établir l'égalité : $E_n(2) = 2^n - (n+1)$.

5. Une relation de récurrence

Soit n un entier naturel non nul. À toute liste $a = (a_1, a_2, \dots, a_{n+1})$ de S_{n+1} , on associe la liste $\varphi_n(a)$ de S_n obtenue en ôtant l'élément $(n+1)$ de la liste a . Par exemple, dans le cas particulier où $n = 5$, si $a = (3, 4, 1, 5, 6, 2)$, alors $\varphi_5(a) = (3, 4, 1, 5, 2)$, et si $a = (6, 3, 4, 1, 5, 2)$ alors $\varphi_5(a) = (3, 4, 1, 5, 2)$.

- (a) Soit $b = (a_1, a_2, \dots, a_n)$ un élément de S_n . Comment s'écrivent les éléments de S_{n+1} dont l'image par φ_n est égale à b ?
- (b) Soit k dans $\llbracket 1, n \rrbracket$ et soit b dans S_n tels que $M(b) = k$. Quelles sont les valeurs possibles de $M(a)$ pour un élément a de S_{n+1} dont l'image par φ_n est égale à b ?
- (c) Établir, pour tout entier k de $\llbracket 1, n \rrbracket$, l'égalité :

$$E_{n+1}(k+1) = (k+1)E_n(k+1) + (n+1-k)E_n(k).$$

Vérifier que cette formule tient également pour $k = 0$ et pour tout entier $k > n$.

- (d) Donner, en détaillant le calcul de $E_5(3)$, les valeurs de $E_n(k)$ pour tous les couples d'entiers (n, k) tels que $1 \leq k \leq n \leq 5$. On consignera les résultats dans un tableau, n étant l'indice de ligne et k l'indice de colonne.

6. La formule de Worpitzky

Établir, pour tout entier $n \in \mathbb{N}^*$ et tout entier $k \in \llbracket 1, n \rrbracket$, l'égalité

$$E_n(k) = \sum_{j=1}^k (-1)^{k-j} \binom{n+1}{k-j} j^n.$$

On raisonnera par récurrence sur l'entier n .

7. Une égalité miraculeuse !

Justifier, pour tout entier naturel n non nul, l'égalité :

$$E_{2n-1}(n) = \frac{2}{\pi} (2n-1)! \int_0^{+\infty} \left(\frac{\sin t}{t} \right)^{2n} dt.$$

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2020

Épreuve à option (A) : Mathématiques

Durée : 3h

Les matrices de Hessenberg sont des matrices « presque triangulaires », qui permettent d'économiser les calculs lors de la mise en œuvre d'algorithmes d'analyse numérique. L'objet du problème est d'examiner le comportement asymptotique des puissances de certaines de ces matrices.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Dans tout l'énoncé, n désigne un nombre entier supérieur ou égal à 3.

- Le coefficient situé dans la i -ème ligne et la j -ème colonne d'une matrice M est noté $M[i, j]$.
- Les matrices élémentaires de $\mathcal{M}_n(\mathbb{R})$, dont tous les coefficients sont nuls sauf un, égal à 1, sont notées $E_{i,j}^{(n)}$. Ainsi a-t-on, pour tout couple $(i, j) \in \llbracket 1, n \rrbracket^2$:

$$\forall (\ell, k) \in \llbracket 1, n \rrbracket^2, \quad E_{i,j}^{(n)}[\ell, k] = \delta_i^\ell \delta_j^k = \begin{cases} 1 & \ell = i \text{ et } k = j \\ 0 & \text{sinon} \end{cases}$$

où δ_i^ℓ (respectivement δ_j^k) vaut 1 si $i = \ell$ (respectivement $j = k$), 0 sinon.

• Pour toute matrice M de $\mathcal{M}_n(\mathbb{R})$, on note u_M (respectivement v_M) l'endomorphisme de $\mathbb{R}^n$ (respectivement $\mathbb{C}^n$) canoniquement associé à M , c'est-à-dire l'endomorphisme dont M est la matrice dans la base canonique $\mathcal{B}_n = (e_1, \dots, e_n)$ de $\mathbb{R}^n$ (respectivement $\mathbb{C}^n$). On appelle *spectre* de M l'ensemble, noté $\text{sp}_{\mathbb{C}}(M)$, des valeurs propres de v_M (il s'agit d'une partie non vide de $\mathbb{C}$).

• Lorsque doit être établie la convergence d'une suite de matrices de $\mathcal{M}_n(\mathbb{R})$, il s'agit de la convergence pour l'une quelconque des normes sur cet espace vectoriel, toutes équivalentes puisqu'il est de dimension finie.

Partie I Matrices de Hessenberg

• On dit qu'une matrice M de $\mathcal{M}_n(\mathbb{R})$ est une *matrice de Hessenberg* si tous ses coefficients $M[i, j]$ situés en dessous de la « sous-diagonale », c'est-à-dire tels que $i \geq j + 2$, sont nuls.

Par exemple, les matrices de Hessenberg de $\mathcal{M}_5(\mathbb{R})$ sont les matrices de la forme

$$\begin{pmatrix} M[1,1] & M[1,2] & M[1,3] & M[1,4] & M[1,5] \\ M[2,1] & M[2,2] & M[2,3] & M[2,4] & M[2,5] \\ 0 & M[3,2] & M[3,3] & M[3,4] & M[3,5] \\ 0 & 0 & M[4,3] & M[4,4] & M[4,5] \\ 0 & 0 & 0 & M[5,4] & M[5,5] \end{pmatrix}.$$

1. Pour tout entier n supérieur ou égal à 3, on note $\mathcal{H}_n(\mathbb{R})$ l'ensemble des matrices de Hessenberg de $\mathcal{M}_n(\mathbb{R})$.
 - a) Justifier que $\mathcal{H}_n(\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$.
 - b) En donner une base et la dimension.
2. Soit a un réel strictement positif. On note $H_+(a)$ et $H_-(a)$ les deux matrices de Hessenberg de $\mathcal{M}_3(\mathbb{R})$ définies par

$$H_+(a) = \begin{pmatrix} a & a & 0 \\ 1 & a & 0 \\ 0 & 0 & a + \sqrt{a} \end{pmatrix} \quad \text{et} \quad H_-(a) = \begin{pmatrix} a & a & a \\ -1 & a & a \\ 0 & 0 & a \end{pmatrix}.$$

a) Trouver les valeurs propres de la matrice $H_+(a)$ et montrer qu'elle est diagonalisable dans $\mathbb{R}$.

b) Soit $r \in \mathbb{R}_+^*$. Démontrer que la suite de matrices $\left(\frac{1}{r^p} (H_+(a))^p \right)_{p \in \mathbb{N}}$ est convergente si, et seulement si, le réel r est supérieur ou égal à $a + \sqrt{a}$.

c) Soit $Q(a) = \lim_{p \rightarrow +\infty} \frac{1}{(a + \sqrt{a})^p} (H_+(a))^p$.

Justifier que l'endomorphisme $u_{Q(a)}$ de $\mathbb{R}^n$ est un projecteur, dont on précisera le rang, l'image et le noyau.

d) Soit $r \in \mathbb{R}_+^*$. Démontrer que la suite de matrices $\left(\frac{1}{r^p} (H_-(a))^p \right)_{p \in \mathbb{N}}$ est convergente si, et seulement si, le réel r est strictement supérieur à $\sqrt{a(1+a)}$.

3. Soit $(a, b) \in \mathbb{R}^2$. On note $J_n = \sum_{j=1}^{n-1} E_{j+1,j}^{(n)}$ et $H_n(a, b) = a I_n + b J_n$.

a) Calculer les puissances successives de la matrice J_3 et donner, sans démonstration, une expression générale des puissances $(J_n)^q$ ($q \in \mathbb{N}$).

b) Établir, pour tout entier $p \geq n - 1$, l'égalité

$$(H_n(a, b))^p = \sum_{j=1}^n \sum_{i=j}^n \binom{p}{i-j} a^{p-i+j} b^{i-j} E_{i,j}^{(n)}$$

et en déduire que, si a et b sont différents de 0, alors tous les coefficients de la matrice $(H_n(a, b))^{n-1}$ situés sur ou en dessous de la diagonale (c'est-à-dire dont l'indice de ligne est supérieur ou égal à l'indice de colonne) sont différents de 0.

c) Pour quelles valeurs de (a, b) la suite $\left((H_n(a, b))^p \right)_{p \in \mathbb{N}}$ est-elle convergente?

Partie II Rayon spectral

Dans cette partie, on considère une matrice M de $\mathcal{M}_n(\mathbb{R})$ et on note $\rho(M)$ le plus grand des modules des éléments du spectre de M , appelé *rayon spectral* de M :

$$\rho(M) = \text{Max} \{ |\lambda| ; \lambda \in \text{sp}_{\mathbb{C}}(M) \} .$$

4. Soit ε un nombre réel strictement positif et P_ε la matrice diagonale de $\mathcal{M}_n(\mathbb{R})$ dont les coefficients diagonaux sont donnés par :

$$\forall i \in \llbracket 1, n \rrbracket, P_\varepsilon[i, i] = \varepsilon^i .$$

a) Calculer, pour toute matrice triangulaire supérieure T de $\mathcal{M}_n(\mathbb{R})$, les coefficients de la matrice $P_\varepsilon^{-1} T P_\varepsilon$.

b) En déduire que, pour tout réel $\alpha > 0$, il existe une matrice triangulaire semblable à T dont tous les coefficients non diagonaux sont de modules inférieurs ou égaux à α .

c) Démontrer que pour tout réel $r > \rho(M)$, il existe une matrice inversible P à coefficients complexes pour laquelle la matrice $M' = P^{-1} M P$ est triangulaire et le module de chacun de ses coefficients au plus égal à r , c'est-à-dire :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, |M'[i, j]| \leq r .$$

5. Justifier, pour tout réel $r > \rho(M)$, la convergence de la suite $\left(\frac{1}{r^p} M^p \right)_{p \in \mathbb{N}}$ vers la matrice nulle.
6. On suppose dans cette question que le polynôme caractéristique de M possède au moins une racine réelle simple, que l'on note λ .

On note u_M^* l'endomorphisme de $\mathbb{R}^n$ de u_M , dont la matrice dans la base canonique est la transposée ${}^t M$ de la matrice M .

- a) Justifier l'existence de deux éléments non nuls x et y de $\mathbb{R}^n$ vérifiant

$$u_M(x) = \lambda x \quad \text{et} \quad u_M^*(y) = \lambda y .$$

b) Démontrer que l'orthogonal $\text{Vect}\{y\}^\perp$ de la droite vectorielle engendrée par le vecteur y , pour le produit scalaire usuel de $\mathbb{R}^n$, est stable par l'endomorphisme u_M et que λ n'est pas une valeur propre de l'endomorphisme de $\text{Vect}\{y\}^\perp$ induit par u_M .

c) En déduire que $\text{Vect}\{y\}^\perp$ et $\text{Vect}\{x\}$ sont deux sous-espaces vectoriels supplémentaires de $\mathbb{R}^n$.

d) On note Q la matrice dans la base canonique du projecteur de $\mathbb{R}^n$ d'image $\text{Vect}\{x\}$ et de noyau $\text{Vect}\{y\}^\perp$.

Démontrer que, si λ est égal au rayon spectral $\rho(M)$ et si tous les éléments de $\text{sp}_{\mathbb{C}}(M)$ distincts de λ ont un module strictement inférieur à $\rho(M)$, alors la suite $\left(\frac{1}{\rho(M)^p} M^p \right)_{p \in \mathbb{N}}$ est convergente et de limite Q .

Partie III Matrices irréductibles

Pour toute partie J de $[[1, n]]$, on note F_J le sous-espace vectoriel de $\mathbb{R}^n$ engendré par les vecteurs e_j de la base canonique de $\mathbb{R}^n$ dont l'indice j appartient à J :

$$F_J = \text{Vect}\{e_j; j \in J\}$$

(qui est réduit au vecteur nul lorsque J est vide).

• On dit qu'une matrice M de $\mathcal{M}_n(\mathbb{R})$ est *J-réduite* si F_J est stable par l'endomorphisme u_M et on note $\mathcal{M}_n(\mathbb{R})[J]$ l'ensemble des matrices *J-réduites* de $\mathcal{M}_n(\mathbb{R})$:

$$\mathcal{M}_n(\mathbb{R})[J] = \{M \in \mathcal{M}_n(\mathbb{R}); u_M(F_J) \subseteq F_J\}.$$

• On dit qu'une matrice M de $\mathcal{M}_n(\mathbb{R})$ est *irréductible* s'il n'existe aucune partie J de $[[1, n]]$, non vide et distincte de $[[1, n]]$, pour laquelle M est *J-réduite*.

7. Exemples

a) Pour quelles parties non vides J de $\{1, 2, 3\}$ les matrices $H_+(a)$ et $H_-(a)$ de la question 2 sont-elles *J-réduites*?

b) Démontrer qu'aucune des matrices $H_n(a, b)$ de la question 3 n'est irréductible.

8. Soit $M \in \mathcal{M}_n(\mathbb{R})$.

Démontrer que, pour que M soit irréductible, il suffit qu'il existe un entier naturel p pour lequel tous les coefficients de la matrice M^p sont différents de 0.

Pour traiter la dernière question du problème, on admettra les deux résultats suivants (théorèmes de Perron-Frobenius).

• Si une matrice $N \in \mathcal{M}_n(\mathbb{R})$ est à coefficients positifs ou nuls, alors $\rho(N)$ appartient au spectre de N .

• Si une matrice $N \in \mathcal{M}_n(\mathbb{R})$ est à coefficients strictement positifs, alors $\rho(N)$ en est la seule valeur propre de module maximal et la dimension du sous-espace propre qui lui est associé est égale à 1.

9. Soit H une matrice de $\mathcal{H}_n(\mathbb{R})$ dont tous les coefficients situés sur et au dessus de la sous-diagonale sont strictement positifs :

$$\forall (i, j) \in [[1, n]]^2, i \leq j + 1 \implies H[i, j] > 0.$$

a) Justifier l'existence de deux réels strictement positifs a et b pour lesquels tous les coefficients de $H - H_n(a, b)$ sont positifs ou nuls.

b) En déduire que H est irréductible.

c) Démontrer que la suite $\left(\frac{1}{\rho(H)^p} H^p\right)_{p \in \mathbb{N}}$ converge vers une matrice qui est canoniquement associée à un projecteur de rang 1.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2020

Épreuve à option (B) : Probabilités

Durée : 3h

L'objet du problème est l'étude de deux mesures de risques fréquemment utilisées en gestion financière : la *Value-at-Risk* (VaR) et la *Tail-VaR*. Les trois parties peuvent être traitées de manière indépendante. La notation tiendra bien évidemment compte de la clarté et de la précision des réponses.

Notations et rappels

Pour l'ensemble du problème, on considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ où $\mathcal{A}$ est une tribu de parties de Ω et $\mathbb{P}$ une mesure de probabilité.

- On note $\mathcal{X}$ l'ensemble des **variables aléatoires réelles**. Un élément de $\mathcal{X}$ est donc une fonction $X : \Omega \rightarrow \mathbb{R}$ telle que pour tout $x \in \mathbb{R}$, $[X \leq x] = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{A}$.
- Soit $X \in \mathcal{X}$. La fonction F définie pour tout $x \in \mathbb{R}$ par $F(x) = \mathbb{P}([X \leq x])$ est la **fonction de répartition** de X . On admettra que cette fonction est nécessairement croissante et continue à droite.
- Un **vecteur aléatoire** de taille $n \geq 1$ est un n -uplet $(X_1, \dots, X_n)$ où pour tout $i \in \{1, \dots, n\}$, $X_i \in \mathcal{X}$ et pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$, $[X_1 \leq x_1] \cap \dots \cap [X_n \leq x_n] \in \mathcal{A}$.
- Deux vecteurs aléatoires $(X_1, \dots, X_n)$ et $(Y_1, \dots, Y_n)$ sont dits **égaux en loi** si pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$, $\mathbb{P}([X_1 \leq x_1] \cap \dots \cap [X_n \leq x_n]) = \mathbb{P}([Y_1 \leq x_1] \cap \dots \cap [Y_n \leq x_n])$. L'égalité en loi est notée $(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (Y_1, \dots, Y_n)$.
- Si la fonction de répartition de $X \in \mathcal{X}$ est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ (éventuellement privé d'un ensemble fini de points) alors X est une **variable aléatoire à densité**. En notant $f = F'$ la dérivée de F on a pour toute fonction continue positive $g : \mathbb{R} \rightarrow [0, \infty[$

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

- Pour tout $X \in \mathcal{X}$ admettant une espérance $\mathbb{E}(X)$ et une variance $\mathbb{V}(X)$, l'**inégalité de Bienaymé-Tchebychev** assure que pour tout $\varepsilon > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

- Soit $(X_n \in \mathcal{X}, n \in \mathbb{N})$ une suite de variables aléatoires. On dit que X_n **converge en probabilité** vers $c \in \mathbb{R}$ ($X_n \xrightarrow{\mathbb{P}} c$) si pour tout $\varepsilon > 0$, $\mathbb{P}(|X_n - c| \geq \varepsilon) \rightarrow 0$ lorsque $n \rightarrow \infty$. Pour toute fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ continue, si $X_n \xrightarrow{\mathbb{P}} c$ alors $g(X_n) \xrightarrow{\mathbb{P}} g(c)$.
- Soit $(X_n \in \mathcal{X}, n \in \mathbb{N})$ une suite de variables aléatoires que l'on supposera indépendantes et de même loi. Si ces variables aléatoires admettent une espérance égale à $\mu \in \mathbb{R}$, la **loi faible des grands nombres** assure que

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} \mu.$$

Introduction du problème

En gestion financière, il est primordial de pouvoir évaluer le risque associé à un portefeuille d'actifs. Plus précisément, si la perte d'un portefeuille sur une période de temps donnée est modélisée par une variable aléatoire X , on souhaite quantifier le risque d'une perte importante de capital en résumant les valeurs prises par X par une unique valeur réelle $\mathcal{R}(X)$ appelée **mesure de risque**. L'objectif de ce problème est l'étude de deux mesures de risque classiques en gestion financière : la **Value-at-Risk (VaR)** et la **Tail-VaR**.

Dans un premier temps, nous supposons que les variables aléatoires $X \in \mathcal{X}$ considérées admettent une fonction de répartition $F(\cdot) = \mathbb{P}([X \leq \cdot])$ vérifiant la propriété suivante : il existe un sous ensemble $I \subset \mathbb{R}$ tel que F est une bijection de I dans $]0, 1[$.

Dans ce cadre¹, la **Value-at-Risk** au niveau $\alpha \in]0, 1[$ d'une variable aléatoire $X \in \mathcal{X}$ modélisant la perte d'un portefeuille d'actifs est donnée par

$$\text{VaR}_\alpha(X) = F^{-1}(1 - \alpha),$$

où F^{-1} est la bijection réciproque de F . Il est évident que la valeur $\text{VaR}_\alpha(X)$ est telle que $\mathbb{P}([X \leq \text{VaR}_\alpha(X)]) = 1 - \alpha$. On a ainsi une confiance de $(1 - \alpha)\%$ que la perte du portefeuille n'excède pas la Value-at-Risk. Ainsi, plus la valeur de $\text{VaR}_\alpha(X)$ est élevée, plus l'investissement dans le portefeuille d'actifs semble risqué. En pratique la valeur de α sera choisie proche de 0 (typiquement $\alpha = 0.01$ ou $\alpha = 0.05$).

La seconde mesure de risque que l'on considère dans ce problème est la **Tail-VaR** au niveau $\alpha \in]0, 1[$ définie pour tout $X \in \mathcal{X}_1 = \{X \in \mathcal{X} \mid \mathbb{E}(|X|) < \infty\} \subset \mathcal{X}$ par

$$\text{TVaR}_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha \text{VaR}_s(X) ds.$$

Nous montrerons dans la partie III que l'intégrale ci-dessus est bien définie pour tout $X \in \mathcal{X}_1$. La Tail-VaR peut donc être vue comme la moyenne des Value-at-Risk de niveau $s \in]0, \alpha]$. Cette mesure de risque présente l'avantage de prendre en compte l'information apportée par les valeurs de X supérieures à la Value-at-Risk de niveau α . Ici encore, une grande valeur de $\text{TVaR}_\alpha(X)$ est associée à un investissement risqué.

I – Application

Vous souhaitez choisir entre deux portefeuilles d'actifs P_1 et P_2 afin d'effectuer un investissement le moins risqué possible. La perte (en K€, sur une période de 10 jours) du portefeuille P_1 est modélisée par $X_1 \in \mathcal{X}$ et celle de P_2 par $X_2 \in \mathcal{X}$. Les fonctions de répartition F_1

1. Le cas des variables aléatoires dont la fonction de répartition est non nécessairement bijective sera traité dans la partie III.

et F_2 de X_1 et X_2 sont données par

$$F_1(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - (1+x)^{-\lambda} & \text{si } x \geq 0 \end{cases} \quad \text{et } F_2(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp(-\lambda x) & \text{si } x \geq 0, \end{cases}$$

où $\lambda > 0$. Pour effectuer votre choix, vous mesurez dans un premier temps le risque par la Value-at-Risk.

- I-1) Calculer en fonction de $\lambda > 0$ et $\alpha \in]0, 1[$ les valeurs $\text{VaR}_\alpha(X_1)$ et $\text{VaR}_\alpha(X_2)$.
- I-2) Montrer que pour tout $x > 1$, $\ln(x) < x - 1$.
- I-3) Montrer que pour tout $\alpha \in]0, 1[$ et $\lambda > 0$, $\text{VaR}_\alpha(X_1) > \text{VaR}_\alpha(X_2)$. Conclure quant au choix du portefeuille.

Vous souhaitez conforter votre choix de portefeuille en mesurant également le risque par la Tail-VaR.

- I-4) Déterminer le sous-ensemble $L \subset]0, \infty[$ tel que pour tout $\lambda \in L$, $\mathbb{E}(|X_1|) = \mathbb{E}(X_1) < \infty$.
- I-5) Montrer que pour tout $\lambda > 0$, $\mathbb{E}(|X_2|) = \mathbb{E}(X_2) < \infty$.
- I-6) Pour tout $\lambda \in L$ et $\alpha \in]0, 1[$, calculer les valeurs $\text{TVaR}_\alpha(X_1)$ et $\text{TVaR}_\alpha(X_2)$.
- I-7) Montrer que pour tout $x > 1$ et $\lambda > 1$, $\lambda/(\lambda - 1)x - 1 > 1/\lambda + \ln(x)$.
- I-8) Pour tout $\lambda \in L$ et $\alpha \in]0, 1[$, comparer les valeurs $\text{TVaR}_\alpha(X_1)$ et $\text{TVaR}_\alpha(X_2)$ et conclure quant au choix du portefeuille.

Dans cette application, les fonctions de répartition F_1 et F_2 sont continues sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R} \setminus \{0\}$ avec de plus $F_1^{-1}(1 - \alpha) \rightarrow \infty$ et $F_2^{-1}(1 - \alpha) \rightarrow \infty$ lorsque $\alpha \rightarrow 0$. Dans ce cas, la Tail-VaR peut être écrite différemment. C'est ce que l'on se propose de montrer dans la question suivante.

- I-9) Dans le cas où la fonction de répartition F de X est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ (éventuellement privé d'un ensemble fini de points) avec de plus $F^{-1}(1 - \alpha) \rightarrow \infty$ lorsque $\alpha \rightarrow 0$, montrer, en effectuant un changement de variable judicieux, que

$$\text{TVaR}_\alpha(X) = \frac{1}{\alpha} \int_{F^{-1}(1-\alpha)}^{\infty} t f(t) dt = \mathbb{E}(X \mid X > F^{-1}(1 - \alpha)),$$

où f est la dérivée de F . Pour la dernière égalité, on rappelle que pour tout $A \in \mathcal{A}$ tel que $\mathbb{P}(A) \neq 0$, $\mathbb{E}(X|A) = \mathbb{E}(X\mathbb{I}_A)/\mathbb{P}(A)$ où $\mathbb{I}_A$ est la fonction indicatrice².

- I-10) Vérifier l'égalité ci-dessus pour les variables aléatoires X_1 et X_2 .

II - Estimation de la Value-at-Risk

Soit $n \geq 1$ un entier et soient $x_1, \dots, x_n$ les pertes (en k€, sur 10 jours) observées pour un portefeuille d'actifs. On admet que ses valeurs sont les réalisations de n variables aléatoires $X_1, \dots, X_n$ supposées indépendantes et de même loi que $X \in \mathcal{X}$ de fonction de répartition F (ainsi, pour tout $x \in \mathbb{R}$, $\mathbb{P}([X_1 \leq x]) = \dots = \mathbb{P}([X_n \leq x]) = \mathbb{P}([X \leq x]) = F(x)$).

2. Pour tout $E \subset \mathbb{R}$, $\mathbb{I}_E : \mathbb{R} \rightarrow \{0, 1\}$ est la fonction telle que $\mathbb{I}_E(x) = 1$ si $x \in E$ et $\mathbb{I}_E(x) = 0$ si $x \notin E$.

On notera $X_{1,n} \leq \dots \leq X_{n,n}$ les statistiques d'ordre associées aux variables aléatoires $X_1, \dots, X_n$. Plus précisément, le vecteur aléatoire $(X_{1,n}, \dots, X_{n,n})$ est tel que pour tout $\omega \in \Omega$,

$$(X_{1,n}(\omega), \dots, X_{n,n}(\omega))$$

est le vecteur obtenu en rangeant par ordre croissant les valeurs $X_1(\omega), \dots, X_n(\omega)$. Les quatre questions suivantes ont pour but de vous familiariser avec la notion de statistiques d'ordre.

II-1) Pour $n = 4$, le tableau suivant donne les valeurs des variables aléatoires $X_1, \dots, X_4$ aux points $\omega_1 \in \Omega$ et $\omega_2 \in \Omega$.

$X_1(\omega_1) = 0.73$	$X_2(\omega_1) = 0.94$	$X_3(\omega_1) = 0.62$	$X_4(\omega_1) = 0.51$
$X_1(\omega_2) = 0.21$	$X_2(\omega_2) = 0.79$	$X_3(\omega_2) = 0.81$	$X_4(\omega_2) = 0.31$

Quelle est la valeur de $X_{3,4}(\omega_1)$? Celle de $X_{3,4}(\omega_2)$?

II-2) Expliquer pourquoi pour tout $x \in \mathbb{R}$ on a $[X_{n,n} \leq x] = [X_1 \leq x] \cap \dots \cap [X_n \leq x]$.

II-3) En déduire que pour tout $x \in \mathbb{R}$, $\mathbb{P}[X_{n,n} \leq x] = F^n(x)$.

II-4) Donner l'expression de $\mathbb{P}[X_{1,n} \leq x]$ en fonction de $F(x)$.

Nous souhaitons estimer la Value-at-Risk au niveau $\alpha \in]0, 1[$, c'est-à-dire trouver une variable aléatoire observable qui converge (dans un sens que l'on précisera plus loin) vers $\text{VaR}_\alpha(X)$. Nous proposons l'estimateur suivant

$$\widehat{\text{VaR}}_{n,\alpha}(X) = X_{[n(1-\alpha)],n},$$

où $[x] = \min\{k \in \mathbb{Z} \mid x \leq k\}$ est la partie entière par excès de x . En notant $x_{1,n} \leq \dots \leq x_{n,n}$ les valeurs observées $x_1, \dots, x_n$ rangées par ordre croissant, la valeur observée de l'estimateur est donc $x_{[n(1-\alpha)],n}$.

L'objectif de cette partie est de montrer que pour tout $\alpha \in]0, 1[$, $\widehat{\text{VaR}}_{n,\alpha}(X) \xrightarrow{\mathbb{P}} \text{VaR}_\alpha(X)$. Pour ce faire, on doit dans un premier temps démontrer un résultat connu sous le nom de *méthode de la transformée inverse*.

On note $U_{1,n} \leq \dots \leq U_{n,n}$ les statistiques d'ordre associées à des variables aléatoires indépendantes $U_1, \dots, U_n$ suivant chacune la loi uniforme sur $[0, 1]$ (i.e. pour tout $i = 1, \dots, n$, U_i est une variable aléatoire à densité de fonction densité $\mathbb{I}_{[0,1]}$).

II-5) *Méthode de la transformée inverse*: Montrer que $X_i \stackrel{\mathcal{L}}{=} F^{-1}(U_i)$ pour tout $i \in \{1, \dots, n\}$.

Autrement dit, montrer que $F(x) = \mathbb{P}([F^{-1}(U_i)] \leq x)$ pour tout $x \in \mathbb{R}$.

Cette égalité en loi permet notamment de simuler n'importe quelle loi de probabilité à partir d'une simulation d'une loi uniforme sur $[0, 1]$.

II-6) En déduire que $(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (F^{-1}(U_1), \dots, F^{-1}(U_n))$.

On admettra pour la suite que l'égalité en loi de la question II-6) implique l'égalité en loi pour les statistiques d'ordre, i.e., que $X_{i,n} \stackrel{\mathcal{L}}{=} F^{-1}(U_{i,n})$ pour tout $i = 1, \dots, n$.

On admettra également que pour tout $i = 1, \dots, n$

$$U_{i,n} \stackrel{\mathcal{L}}{=} \sum_{j=1}^i E_j \bigg/ \sum_{j=1}^{n+1} E_j,$$

où $E_1, \dots, E_{n+1}$ sont des variables aléatoires indépendantes suivant une loi exponentielle standard c'est-à-dire telles que pour tout $i = 1, \dots, n+1$,

$$\mathbb{P}([E_i \leq x]) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp(-x) & \text{si } x \geq 0. \end{cases}$$

II-7) Calculer l'espérance d'une loi exponentielle standard.

II-8) Quelle suite $(u_n)_{n \in \mathbb{N}^*}$ doit-on prendre pour que $U_{[n(1-\alpha)],n} \stackrel{\mathcal{L}}{=} u_n S_n / T_n$ avec

$$S_n = \frac{1}{[n(1-\alpha)]} \sum_{j=1}^{[n(1-\alpha)]} E_j \text{ et } T_n = \frac{1}{n+1} \sum_{j=1}^{n+1} E_j$$

II-9) Montrer que $[n(1-\alpha)] \rightarrow \infty$ et $u_n \rightarrow 1-\alpha$ lorsque $n \rightarrow \infty$. Pour ce faire, vous pourrez utiliser le fait que $[n(1-\alpha)] - n(1-\alpha) \in [0, 1[$.

II-10) En utilisant un résultat donné en début d'énoncé ainsi que les réponses aux questions précédentes, montrer que $S_n \xrightarrow{\mathbb{P}} 1$ et $T_n \xrightarrow{\mathbb{P}} 1$.

On déduit facilement des questions II-8) à II-10) que $U_{[n(1-\alpha)],n} \xrightarrow{\mathbb{P}} 1-\alpha$.

II-11) En déduire que $\widehat{\text{VaR}}_{n,\alpha}(X) \xrightarrow{\mathbb{P}} \text{VaR}_\alpha(X)$.

Partie III – Value-at-Risk et Tail-VaR

A ce stade, nous avons défini les deux mesures de risque (Value-at-Risk et Tail-VaR) uniquement pour des variables aléatoires $X \in \mathcal{X}$ dont la fonction de répartition $F(\cdot) = \mathbb{P}([X \leq \cdot])$ est bijective (de $I \subset \mathbb{R}$ dans $]0, 1[$). Nous allons dans cette partie généraliser ces définitions au cas d'une fonction de répartition quelconque. La fonction F n'étant plus nécessairement bijective, la définition de la Value-at-Risk utilisée jusqu'à présent n'a plus de sens, la bijection réciproque F^{-1} n'existant pas forcément. On propose alors de remplacer dans la définition de la Value-at-Risk F^{-1} par l'inverse généralisée $F^{\leftarrow}$ de F . Plus précisément, pour une variable aléatoire $X \in \mathcal{X}$ de fonction de répartition F quelconque, la Value-at-Risk au niveau $\alpha \in]0, 1[$ est donnée par

$$\text{VaR}_\alpha(X) = F^{\leftarrow}(1-\alpha) = \inf\{x \in \mathbb{R} \mid F(x) \geq 1-\alpha\}.$$

On remarque facilement que $\mathbb{P}([X \leq \text{VaR}_\alpha(X)]) \geq 1-\alpha$. On a ainsi une confiance supérieure à $(1-\alpha)\%$ que la perte du portefeuille n'excède pas la valeur $\text{VaR}_\alpha(X)$.

L'objectif des questions suivantes est de démontrer quelques propriétés de la Value-at-Risk.

III-1) Montrer que si F est bijective de I dans $]0, 1[$ alors $F^{\leftarrow}(1-\alpha) = F^{-1}(1-\alpha)$.

III-2) Pour tout $m \in \mathbb{R}$, $\lambda > 0$ et $X \in \mathcal{X}$, montrer que $\text{VaR}_\alpha(\lambda X + m) = \lambda \text{VaR}_\alpha(X) + m$ pour tout $\alpha \in]0, 1[$.

On rappelle que la Tail-VaR au niveau α est définie pour tout $X \in \mathcal{X}_1 = \{X \in \mathcal{X} \mid \mathbb{E}(|X|) < \infty\}$ par

$$\text{TVaR}_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha \text{VaR}_s(X) ds.$$

L'objectif des questions III-3) à III-7) est de s'assurer que la Tail-VaR est bien défini pour tout $X \in \mathcal{X}_1$ c'est-à-dire que pour tout $\alpha \in]0, 1[$ et $X \in \mathcal{X}_1$,

$$\int_0^\alpha |\text{VaR}_s(X)| ds < \infty. \quad (1)$$

III-3) Montrer que pour tout $x \in \mathbb{R}$ et $y \in]0, 1[$ tels que $F^\leftarrow(y) \leq x$ on a $y \leq F(x)$.

III-4) Réciproquement, montrer que pour tout $x \in \mathbb{R}$ et $y \in]0, 1[$ tels que $y \leq F(x)$ on a $F^\leftarrow(y) \leq x$. En déduire que $\{y \in \mathbb{R} \mid F^\leftarrow(y) \leq x\} = \{y \in \mathbb{R} \mid y \leq F(x)\}$.

III-5) Soit U est une variable aléatoire de loi uniforme sur $[0, 1]$. Utiliser le résultat de la question précédente pour montrer que pour tout $x \in \mathbb{R}$, $[F^\leftarrow(U) \leq x] = [U \leq F(x)]$. En déduire que $F^\leftarrow(U) \stackrel{\mathcal{L}}{=} X$.

III-6) Montrer que pour tout $\alpha \in]0, 1[$,

$$\int_0^\alpha |\text{VaR}_s(X)| ds \leq \mathbb{E}(|F^\leftarrow(U)|),$$

où U est une variable aléatoire de loi uniforme sur $[0, 1]$.

III-7) En utilisant l'égalité en loi $F^\leftarrow(U) \stackrel{\mathcal{L}}{=} X$, montrer que l'assertion (1) est vraie pour tout $X \in \mathcal{X}_1$ (aide : si $X \stackrel{\mathcal{L}}{=} Y$, $\mathbb{E}(|X|) = \mathbb{E}(|Y|)$).

III-8) En utilisant les résultats des questions III-2-c) et III-3), montrer que si $Y \leq Z$ alors $\text{TVaR}_\alpha(Y) \leq \text{TVaR}_\alpha(Z)$.

Montrer également que pour tout $m \in \mathbb{R}$ et $\lambda > 0$, $\text{TVaR}_\alpha(\lambda X + m) = \lambda \text{TVaR}_\alpha(X) + m$.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2021

Épreuve de mathématiques

Durée : 4h

L'épreuve est constituée de quatre exercices indépendants.

Exercice 1. Une caractérisation de la loi géométrique

Soit X et Y deux variables aléatoires à valeurs dans $\mathbb{N}^*$, **indépendantes et de même loi**, toutes deux définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$. On pose, pour tout $\omega \in \Omega$,

$$I(\omega) = \min(X(\omega), Y(\omega)), \quad M(\omega) = \max(X(\omega), Y(\omega)) \quad \text{et} \quad D(\omega) = M(\omega) - I(\omega).$$

1. Montrer que I et M sont des variables aléatoires sur $(\Omega, \mathcal{F})$.
2. Dans cette question, on suppose que la loi commune de X et Y est géométrique de paramètre $p \in]0, 1[$. On pourra poser $q = 1 - p$.
 - (a) Reconnaître la loi de la variable I .
 - (b) Calculer, pour tout $(i, d) \in \mathbb{N} \times \mathbb{N}^*$, la probabilité $\mathbf{P}([I = i] \cap [D = d])$ (qu'on pourra noter $\mathbf{P}(I = i, D = d)$).
On séparera les cas $d = 0$ et $d > 0$.
 - (c) Déterminer la loi de la variable D .
 - (d) Vérifier que les variables I et D sont indépendantes.
3. Dans cette question, la loi commune de X et Y est inconnue et on suppose que les variables I et D sont indépendantes.

On note $b := \mathbf{P}(D = 0)$ et, pour tout entier naturel k non nul, $p_k = \mathbf{P}(X = k)$.

On suppose $p_k > 0$ pour tout $k \in \mathbb{N}^*$.

- (a) Exprimer le réel b à l'aide de la famille $(p_i)_{i \in \mathbb{N}^*}$.
- (b) Exprimer, pour tout entier naturel k , la probabilité $\mathbf{P}(I > k)$ à l'aide de la famille $(p_i)_{i \in \mathbb{N}^*}$.
- (c) Soit $k \in \mathbb{N}$. En calculant la probabilité $\mathbf{P}(I > k, D = 0)$ établir l'égalité

$$\sum_{i=k+1}^{+\infty} p_i^2 = b \left(\sum_{i=k+1}^{+\infty} p_i \right)^2.$$

- (d)
 - i. En déduire, pour tout entier naturel k non nul, l'égalité : $(1 - b)p_k = 2b\mathbf{P}(X > k)$.
 - ii. Calculer p_1 en fonction de b puis établir, pour tout entier naturel k non nul, l'égalité : $p_{k+1} = \frac{1-b}{1+b} p_k$.
- (e) En déduire que la loi commune des variables X et Y est géométrique de paramètre p_1 .

Exercice 2. Un calcul de $\zeta(2)$

On pose, pour tout entier naturel n , $C_n = \int_0^{\frac{\pi}{2}} (\cos x)^{2n} dx$ et $D_n = \int_0^{\frac{\pi}{2}} x^2 (\cos x)^{2n} dx$.

1. Établir, pour tout entier naturel n non nul, l'égalité : $C_n = (2n - 1)(C_{n-1} - C_n)$.

On pourra écrire $\cos^{2n} x = \cos x \cos^{2n-1} x$.

2. Établir, pour tout entier naturel n non nul, les égalités

$$\int_0^{\frac{\pi}{2}} (\sin x)^2 (\cos x)^{2n-2} dx = \frac{C_n}{2n-1} = \frac{C_{n-1}}{2n}.$$

3. Établir, pour tout entier naturel n non nul, l'égalité : $C_n = (2n-1)nD_{n-1} - 2n^2D_n$.

4. En déduire, pour tout entier naturel n non nul, l'égalité : $\frac{1}{n^2} = 2\left(\frac{D_{n-1}}{C_{n-1}} - \frac{D_n}{C_n}\right)$.

5. (a) Justifier, pour tout réel $x \in \left[0, \frac{\pi}{2}\right]$, la minoration : $\sin x \geq \frac{2}{\pi}x$.

(b) En déduire, pour tout entier naturel n , la majoration : $D_n \leq \frac{\pi^2}{4} \frac{C_n}{2n+2}$.

6. Prouver l'égalité : $\sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Exercice 3. Trois preuves d'une égalité combinatoire

1. Avec la fonction Beta

On pose, pour tout couple (n, p) d'entiers naturels, $B(n, p) = \int_0^1 t^n (1-t)^p dt$.

(a) Relier, pour tout couple $(n, p) \in (\mathbb{N}^*)^2$, les réels $B(n, p)$ et $B(n+1, p-1)$. En déduire, pour tout couple $(n, p) \in \mathbb{N}^2$, l'égalité : $B(n, p) = \frac{n!p!}{(n+p+1)!}$.

(b) Soit $(m, n) \in \mathbb{N}^* \times \mathbb{N}$. Établir l'égalité

$$(\mathcal{E}) \quad \frac{(m-1)!n!}{(m+n)!} = \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{m+k}.$$

2. Avec des différences finies

On note Δ l'application qui, à chaque suite réelle $u := (u_p)_{p \in \mathbb{N}^*}$, associe la suite

$$\Delta(u) := (u_{p+1} - u_p)_{p \in \mathbb{N}^*}.$$

On note, $\Delta^2 = \Delta \circ \Delta$, $\Delta^3 = \Delta \circ \Delta \circ \Delta$, etc., et $\Delta^0 = \text{Id}_{\mathbb{R}^{\mathbb{N}^*}}$.

(a) Prouver, pour toute suite $u := (u_p)_{p \in \mathbb{N}^*}$, pour tout entier naturel n et tout entier $p \in \mathbb{N}^*$, l'égalité

$$(\Delta^n(u))_p = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} u_{p+k}.$$

(b) En utilisant la suite $u := \left(\frac{1}{p}\right)_{p \in \mathbb{N}^*}$ établir à nouveau l'égalité $(\mathcal{E})$.

3. Par un raisonnement probabiliste

On rappelle que si $n \in \mathbb{N}^*$ et $A_1, A_2, \dots, A_n$ sont des événements d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ on dispose de la formule du crible suivante

$$\mathbf{P}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}).$$

- (a) Soit $A_1, A_2, \dots, A_n$ et B des événements d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$. Établir l'égalité

$$\mathbf{P}(B \cap \bigcap_{i=1}^n A_i) = \mathbf{P}(B) + \sum_{k=1}^n (-1)^k \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbf{P}(B \cap A_{i_1}^c \cap A_{i_2}^c \cap \dots \cap A_{i_k}^c),$$

(où l'on note E^c le complémentaire d'un événement E).

Soit $(m, n) \in (\mathbb{N}^*)^2$. On considère une urne contenant une boule noire notée N , n boules blanches notées $B_1, B_2, \dots, B_n$ et $m - 1$ boules rouges notées $R_1, R_2, \dots, R_{m-1}$. On effectue des tirages successifs d'une boule dans cette urne, sans remise de la boule dans l'urne après tirage, jusqu'à épuisement de l'urne. Un résultat de cette expérience aléatoire est donc une $(m+n)$ -liste (ordonnée ...) composée des $m+n$ symboles $B_1, B_2, \dots, B_n, R_1, R_2, \dots, R_{m-1}$ et N . On note B l'événement constitué des tirages où la boule noire est tirée avant chacune des boules rouges, et, pour tout entier $i \in \llbracket 1, n \rrbracket$, A_i l'événement constitué des tirages où la boule noire est tirée après la boule blanche B_i .

- (b) Calculer la probabilité $\mathbf{P}(B)$.

- (c) Calculer la probabilité $\mathbf{P}(B \cap \bigcap_{i=1}^n A_i)$.

- (d) Prouver, pour tout entier $k \in \llbracket 1, n \rrbracket$ et tout k -liste $(i_1, i_2, \dots, i_k)$ avec $1 \leq i_1 < i_2 < \dots < i_k \leq n$, l'égalité

$$\mathbf{P}(B \cap A_{i_1}^c \cap A_{i_2}^c \cap \dots \cap A_{i_k}^c) = \frac{1}{m+k}.$$

- (e) i. Soit $k \in \llbracket 1, n \rrbracket$. Déterminer – en le justifiant – le nombre de k -listes $(i_1, i_2, \dots, i_k)$ d'entiers telles que $1 \leq i_1 < i_2 < \dots < i_k \leq n$.
ii. En déduire à nouveau l'égalité (e).

Exercice 4. Une propriété des matrices symétriques

On note, pour tout entier naturel n non nul, $\mathbb{M}_{n,1}(\mathbb{R})$ l'ensemble des matrices réelles colonnes de taille n que l'on munit de sa structure euclidienne standard où le produit de deux matrices colonnes X et Y est tXY . On note $S_n(\mathbb{R})$ l'ensemble des matrices réelles carrées d'ordre n et symétriques.

On note, pour tout entier naturel n non nul et pour toute matrice $A \in S_n(\mathbb{R})$,

$$R(A) = \{ {}^tXAX; {}^tXX = 1 \}.$$

On dit qu'une matrice $A \in \mathbb{M}_n(\mathbb{R})$ est **orthogonalement semblable** à une matrice $B \in \mathbb{M}_n(\mathbb{R})$ s'il existe une matrice $P \in O_n(\mathbb{R})$ (i.e. vérifiant l'égalité ${}^tPP = I_n$) pour laquelle $B = {}^tPAP$.

- Soit $(A, B) \in S_n(\mathbb{R})^2$. On suppose que A est orthogonalement semblable à B .
Prouver l'égalité : $R(A) = R(B)$.
- Soit $n \in \mathbb{N}^*$ et $A \in S_n(\mathbb{R})$. On note $\lambda_1, \lambda_2, \dots, \lambda_n$ les valeurs propres de A rangées dans l'ordre croissant i.e. $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.
(a) Prouver l'inclusion : $R(A) \subset [\lambda_1, \lambda_n]$.

(b) Établir l'égalité : $R(A) = [\lambda_1, \lambda_n]$.

(c) Montrer que si la matrice A est de trace nulle alors 0 est élément de $R(A)$. Dans le cas où $n = 3$, la réciproque est-elle vraie ?

On appelle diagonale d'une matrice $M := (m_{i,j})_{1 \leq i,j \leq n}$ la liste $(m_{1,1}, m_{2,2}, \dots, m_{n,n})$ de ses éléments diagonaux.

2. Soit $n \in \mathbb{N}^*$ et $A \in S_n(\mathbb{R})$. On suppose qu'on dispose d'une matrice $P \in O_n(\mathbb{R})$ pour laquelle tPAP **a pour diagonale** $(\text{Tr } A, 0, \dots, 0)$ (où $\text{Tr } A$ désigne la trace de A). Vérifier que $\text{Tr } A$ est élément de $R(A)$.

3. (a) Donner un exemple de matrice $A \in S_2(\mathbb{R})$ dont la trace n'est pas élément de $R(A)$.

(b) Soit $A \in S_2(\mathbb{R})$. Montrer que $\text{Tr } A \in R(A)$ si et seulement si $0 \in R(A)$.

4. Soit $A \in S_2(\mathbb{R})$. On suppose que $\text{Tr } A \in R(A)$. Montrer que A est orthogonalement semblable à une matrice dont la diagonale est $(\text{Tr } A, 0)$.

5. Soit n un entier avec $n \geq 2$. On suppose que toute matrice $A \in S_n(\mathbb{R})$ telle que $\text{Tr } A \in R(A)$ est orthogonalement semblable à une matrice ayant pour diagonale $(\text{Tr } A, 0, \dots, 0)$.

Soit $A \in S_{n+1}(\mathbb{R})$ telle que $\text{Tr } A \in R(A)$.

(a) Justifier l'existence d'une matrice colonne $C \in \mathbb{M}_{n,1}(\mathbb{R})$, d'une matrice ligne $L \in \mathbb{M}_{1,n}(\mathbb{R})$ et d'une matrice $B \in S_n(\mathbb{R})$ pour lesquelles la matrice A est orthogonalement semblable à la matrice par blocs $\begin{pmatrix} \text{Tr } A & L \\ C & B \end{pmatrix}$.

(b) Que vaut $\text{Tr } B$? En déduire que $\text{Tr } B \in R(B)$.

(c) Conclure que la matrice A est orthogonalement semblable à une matrice de diagonale $(\text{Tr } A, 0, \dots, 0)$.

6. Soit $A \in S_n(\mathbb{R})$. Montrer qu'il existe un réel a pour lequel la matrice A est orthogonalement semblable à une matrice dont tous les coefficients diagonaux sont égaux à a .

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2021

Épreuve à option (A) : Mathématiques

Durée : 4h

Le thème du problème est le comportement asymptotique des restes des séries numériques convergentes, à travers l'étude d'exemples variés.

L'énoncé est divisé en quatre parties largement indépendantes, que les candidats ne sont pas tenus de traiter dans l'ordre.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Pour toute suite réelle $u = (u_n)_{n \geq 0}$, on notera $\sum_{n \geq 0} u_n$ la série de terme général u_n et $\sum_{n=0}^{\infty} u_n$ la somme de cette série lorsqu'elle est convergente.

Partie I Exemples de calcul explicite du reste

1. Rappeler pourquoi, lorsque la série $\sum_{n \geq 0} u_n$ est convergente, la suite de terme général

$$\sum_{k=n}^{\infty} u_k \text{ est convergente. Quelle est alors sa limite?}$$

2. Dans cette question, x désigne un nombre réel non nul, de signe quelconque.

- a) Démontrer que la série $\sum_{n \geq 0} \frac{x^{2n}}{(2n)!}$ est convergente. Quelle est sa somme?

- b) Établir, pour tout nombre entier strictement positif n , l'égalité :

$$\sum_{k=n}^{\infty} \frac{x^{2k}}{(2k)!} = \int_0^x \frac{(x-t)^{2n-2}}{(2n-2)!} \operatorname{sh}(t) dt \quad .$$

- c) Donner une expression similaire de $\sum_{k=n}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$, sous la forme d'une intégrale.

3. a) Démontrer que la série de terme général $a_n = \arctan\left(\frac{2n}{n^4 + n^2 + 2}\right)$ est convergente.

- b) Trouver un couple (P, Q) de polynômes de $\mathbb{R}[X]$ qui vérifient :

$$\begin{cases} P(X) - Q(X) = 2X \\ P(X) Q(X) = X^4 + X^2 + 1 \end{cases} \quad .$$

- c) Établir que, pour tout couple (x, y) de nombres réels positifs ou nuls, on a :

$$\arctan\left(\frac{x-y}{1+xy}\right) = \arctan(x) - \arctan(y) \quad .$$

- d) Dédurre des deux questions précédentes que, pour tout entier naturel n , on a :

$$\sum_{k=n}^{\infty} a_k = \frac{\pi}{2} - \arctan(n^2 - n + 1) \quad .$$

- e) La série $\sum_{n \geq 0} \left(\sum_{k=n}^{\infty} a_k \right)$ est-elle convergente?

Partie II Exemples d'évaluation asymptotique du reste

4. Soit $x \in \mathbb{R}$.

Démontrer que la série $\sum_{n \geq 1} \frac{\ln n}{n^x}$ est convergente si, et seulement si, x est strictement supérieur à 1.

5. Dans cette question on suppose que le réel x est strictement supérieur à 1 et, pour tout entier n strictement positif, on note $r_n = \sum_{k=n}^{\infty} \frac{\ln k}{k^x}$.

a) Pour tout réel a strictement positif, justifier l'égalité :

$$\int_a^{+\infty} \frac{\ln t}{t^x} dt = \frac{a^{1-x}(1 + (x-1)\ln a)}{(x-1)^2}.$$

b) Établir, pour tout entier n supérieur ou égal à 4, la double inégalité :

$$\int_n^{+\infty} \frac{\ln t}{t^x} dt \leq r_n \leq \int_{n-1}^{+\infty} \frac{\ln t}{t^x} dt.$$

c) En déduire que r_n est équivalent à $\frac{\ln n}{(x-1)n^{x-1}}$ quand n tend vers l'infini.

6. Sommation des relations de comparaison

On considère deux suites $v = (v_n)_{n \geq 0}$ et $w = (w_n)_{n \geq 0}$ à termes réels non nuls.

On suppose que v_n est équivalent à w_n quand n tend vers l'infini et que la série $\sum_{n \geq 0} v_n$ est convergente.

On rappelle que, si tous les termes de la suite v sont positifs, alors :

- la série $\sum_{n \geq 0} w_n$ est convergente.
- $\sum_{k=n}^{\infty} w_k$ est équivalent à $\sum_{k=n}^{\infty} v_k$ quand n tend vers l'infini.

a) À l'aide d'un contre-exemple, montrer que la première de ces deux propriétés de w ne serait pas assurée si le signe des termes de la suite v n'était pas constant.

b) Montrer de même que, lorsque le signe des termes de la suite v n'est pas constant, il est possible que la série $\sum_{n \geq 0} w_n$ soit convergente mais que $\sum_{k=n}^{\infty} w_k$ ne soit pas équivalent à $\sum_{k=n}^{\infty} v_k$ quand n tend vers l'infini. On pourra utiliser pour w la suite de terme

général $\frac{(-1)^n}{n\sqrt{n}} + \frac{1}{n(n+1)}$.

Pour traiter la question qui suit, on pourra admettre la propriété suivante qui complète les propriétés rappelées précédemment.

Si v et w sont deux suites à termes positifs telles que $v_n = o(w_n)$ (v_n négligeable par rapport à w_n quand n tend vers l'infini) et si les séries $\sum_{n \geq 0} v_n$ et $\sum_{n \geq 0} w_n$ sont convergentes, alors

$$\sum_{k=n}^{\infty} v_k = o\left(\sum_{k=n}^{\infty} w_k\right).$$

7. Un exemple probabiliste

Dans cette question, on considère deux variables aléatoires X et Y définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$.

On suppose que X suit la loi géométrique de paramètre $p \in]0, 1[$, que Y suit la loi de Poisson de paramètre $\lambda (> 0)$ et que les variables aléatoires X et Y sont indépendantes.

a) Pour tout $n \in \mathbb{N}$, calculer $\mathbf{P}(X \geq n)$.

b) Établir, pour tout $n \in \mathbb{N}^*$, l'égalité

$$\mathbf{P}(X + Y = n) = p(1-p)^{n-1} e^{-\lambda} \sum_{k=0}^{n-1} \frac{1}{k!} \left(\frac{\lambda}{1-p} \right)^k.$$

c) En déduire que $\mathbf{P}(Y \geq n)$ est négligeable devant $\mathbf{P}(X + Y \geq n)$ quand n tend vers l'infini.

d) Démontrer que $\mathbf{P}(X + Y \geq n)$ est équivalent à $e^{\lambda p/(1-p)} (1-p)^{n-1}$ quand n tend vers l'infini.

e) En déduire que la série $\sum_{n \geq 0} \mathbf{P}(X + Y \geq n)$ est convergente. Quelle est sa somme?

Partie III Étude du reste comme opérateur

On note L_∞ le $\mathbb{R}$ -espace vectoriel des suites réelles bornées et $\| \cdot \|_\infty$ la norme sur L_∞ définie par

$$\forall u = (u_n)_{n \geq 0}, \quad \|u\|_\infty = \sup \{|u_n|; n \in \mathbb{N}\}.$$

8. Démontrer que l'espace vectoriel F des suites réelles convergentes et de limite nulle est une partie fermée de l'espace vectoriel normé $(L_\infty, \| \cdot \|_\infty)$.

9. On note E l'ensemble des suites réelles $u = (u_n)_{n \geq 0}$ telles que la série $\sum_{n \geq 0} u_n$ est convergente.

a) Justifier que E est un sous-espace vectoriel de F .

b) L'ensemble E est-il une partie fermée de l'espace vectoriel normé $(L_\infty, \| \cdot \|_\infty)$? Quelle est son adhérence?

10. On note Φ l'application qui, à tout élément $u = (u_n)_{n \geq 0}$ de E associe la suite $r = (r_n)_{n \geq 0}$ définie par

$$\forall n \in \mathbb{N}, \quad r_n = \sum_{k=n}^{\infty} u_k$$

a) Calculer $\|u\|_\infty$ et $\|\Phi(u)\|_\infty$ lorsque u est une suite géométrique convergente de premier terme u_0 égal à 1.

b) Démontrer que Φ est une application linéaire et injective de E dans L_∞ , dont l'image est F .

c) On note Ψ la restriction de Φ à E , considérée comme une bijection de E sur F .

On munit E et F de la norme $\|\cdot\|_\infty$.

i) L'application Ψ est-elle continue?

ii) L'application réciproque Ψ^{-1} est-elle continue?

Partie IV Restes de séries alternées

Dans cette partie, f désigne une fonction dérivable sur $[0, +\infty[$, décroissante et convexe, à valeurs strictement positives et de limite nulle en $+\infty$.

On rappelle que, si $v = (v_n)_{n \geq 0}$ est une suite à termes réels strictement positifs, décroissante et de limite nulle, alors :

- la série $\sum_{n \geq 0} (-1)^n v_n$ est convergente
- pour tout $n \in \mathbb{N}$, la somme $\sum_{k=n}^{\infty} (-1)^k v_k$ est positive si n est pair, négative si n est impair, et vérifie l'inégalité

$$\left| \sum_{k=n}^{\infty} (-1)^k v_k \right| \leq v_n.$$

11. Établir, pour tout réel positif ou nul t , la double inégalité :

$$0 \leq \frac{f(t) - f(t+1)}{f(t)} \leq -\frac{f'(t)}{f(t)}.$$

12. Pour tout entier positif ou nul n , on pose : $u_n = (-1)^n f(n)$.

a) Justifier la convergence de la série $\sum_{n \geq 0} u_n$.

b) Pour tout entier positif ou nul n , on pose : $r_n = \sum_{k=n}^{\infty} u_k$.

Démontrer que, pour tout entier positif ou nul n , on a :

i) $|r_n| = \sum_{p=0}^{\infty} (-1)^p f(n+p)$

ii) $0 \leq |r_n| - |r_{n+1}| \leq f(n) - f(n+1).$

c) Montrer que la série $\sum_{n \geq 0} r_n$ est convergente.

d) Démontrer que, si le quotient $\frac{f'(t)}{f(t)}$ tend vers 0 quand le réel t tend vers l'infini alors r_n est équivalent à $\frac{u_n}{2}$ quand n tend vers l'infini.

13. a) Pour quelles valeurs du réel x la série $\sum_{n \geq 1} \frac{(-1)^n \ln n}{n^x}$ est-elle convergente?
- b) Pour ces valeurs, déduire des résultats précédents un équivalent de $\sum_{k=n}^{\infty} \frac{(-1)^k \ln k}{k^x}$ quand n tend vers l'infini.
14. a) Démontrer que la série $\sum_{n \geq 2} \frac{(-1)^n}{n + (-1)^n}$ est convergente.
- b) La série de terme général $\sum_{k=n}^{\infty} \frac{(-1)^k}{k + (-1)^k}$ est-elle convergente?

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2021

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet est composé de quatre problèmes qui peuvent être traités indépendamment. La notation tiendra largement compte de la clarté et de la précision des réponses.

Problème I – Le facteur de Bayes

Le résultat d'une expérience aléatoire est un réel r qui est vu comme la réalisation d'une variable aléatoire R . Pour une même expérience, la loi de la variable aléatoire R peut être définie de différentes manières. Choisir une loi pour R équivaut à choisir un modèle pour l'expérience aléatoire.

Face à ces différentes possibilités de modélisation, une question naturelle est de savoir *a posteriori*, c'est-à-dire après avoir pris connaissance du résultat de l'expérience, si un modèle A est préférable à un autre modèle B . Pour répondre à cette question, on peut calculer le facteur de Bayes défini par

$$f_{A,B} := \log_{10} p_A - \log_{10} p_B,$$

où p_A est la probabilité que $R = r$ sous le modèle A et p_B est la probabilité que $R = r$ sous le modèle B . La notation $\log_{10}$ correspond au logarithme décimal (base 10).

Evidemment, une valeur positive de $f_{A,B}$ suggère que le modèle A est préférable à B . Plus précisément, l'interprétation de la valeur $f_{A,B}$ est donnée dans le tableau suivant.

$f_{A,B}$	Décision
$< -1/2$	Le modèle B est préféré au modèle A .
de $-1/2$ à $1/2$	On ne peut pas distinguer les deux modèles.
$> 1/2$	Le modèle A est préféré au modèle B .

TABLE 1 – Règle de décision du facteur de Bayes

L'objectif de ce problème est d'appliquer cette méthode de décision sur une expérience simple : le "pile ou face".

Questions préliminaires

Soient X_1 et X_2 deux variables aléatoires indépendantes distribuées selon une loi de Bernoulli de paramètres respectifs $q_1 \in]0, 1[$ et $q_2 \in]0, 1[$. On a donc $\mathbb{P}([X_i = 1]) = q_i$ et $\mathbb{P}([X_i = 0]) = 1 - q_i$ pour $i \in \{1, 2\}$.

- I.1) On introduit la variable aléatoire $S_2 = X_1 + X_2$. Déterminer l'ensemble E_2 des valeurs prises par la variable aléatoire S_2 .
- I.2) Pour tout $k \in E_2$, donner l'expression (en fonction de q_1 et q_2) la valeur de la probabilité $\mathbb{P}([S_2 = k])$.
- I.3) Calculer l'espérance et la variance de S_2 .

On considère à présent $n > 2$ variables aléatoires indépendantes $X_1, \dots, X_n$ de loi de Bernoulli avec pour tout $i \in \{1, \dots, n\}$, $\mathbb{P}([X_i = 1]) = q_i \in]0, 1[$. On pose $S_n = X_1 + \dots + X_n$.

I.4) En fonction de $q_1, \dots, q_n$, donner les expressions de $\mathbb{P}([S_n = 0])$ et $\mathbb{P}([S_n = 1])$.

Pour tout $k \in \{1, \dots, n\}$, on note $\mathcal{S}_k$ l'ensemble des vecteurs $(i_1, \dots, i_k) \in \{1, \dots, n\}^k$ tels que $1 \leq i_1 < \dots < i_k \leq n$.

I.5) Donner en fonction de k et n le cardinal de l'ensemble $\mathcal{S}_k$.

I.6) Montrer que

$$\mathbb{P}([S_n = 2]) = \sum_{(i_1, i_2) \in \mathcal{S}_2} q_{i_1} q_{i_2} \prod_{i_\ell \neq i_1, i_2} (1 - q_{i_\ell}).$$

I.7) Pour tout $k \in \{2, \dots, n\}$, donner l'expression de $\mathbb{P}([S_n = k])$.

I.8) Dans le cas où $q_1 = \dots = q_n$, simplifier l'expression de $\mathbb{P}([S_n = k])$ obtenue à la question précédente. Quel nom donne-t-on à la loi de S_n dans ce cas?

Le “pile ou face”

Prenons l'exemple de l'expérience aléatoire consistant à lancer n fois une pièce de monnaie et à compter le nombre de fois où l'on obtient le côté “pile”. Le résultat r de cette expérience est la réalisation de la variable aléatoire R donnée par

$$R = S_n := \sum_{i=1}^n X_i,$$

les variables aléatoires $X_1, \dots, X_n$ étant indépendantes et telles que $X_i = 1$ si le côté “pile” est obtenu au i -ème lancer et $X_i = 0$ sinon.

On compare les modélisations suivantes.

- Modèle $M_{1/2}$: la pièce est équilibrée et donc $\mathbb{P}([X_i = 1]) = 1/2$ pour tout $i \in \{1, \dots, n\}$.
- Modèle M_q : la probabilité d'obtenir le côté “pile” est égale à $q \in]0, 1[$.

I.9) Donner l'expression de la probabilité $p_{1/2}(n, r) := \mathbb{P}([S_n = r])$ lorsque la pièce est supposée équilibrée (modèle $M_{1/2}$).

I.10) Pour tout $r \in \{0, \dots, n\}$, donner l'expression de la probabilité $p_q(n, r) := \mathbb{P}([S_n = r])$ lorsque l'on se place sous le modèle M_q . En déduire l'expression du facteur de Bayes

$$f_{1/2, q}(n, r) = \log_{10} p_{1/2}(n, r) - \log_{10} p_q(n, r).$$

I.11) Dresser le tableau de variation de la fonction $q \mapsto f_{1/2, q}(n, r)$.

On introduit l'ensemble

$$\mathcal{Q} = \left\{ q \in]0, 1[\mid -\frac{1}{2} \leq f_{1/2, q}(n, r) \leq \frac{1}{2} \right\},$$

- I.12) Pour tout $q \in \mathcal{Q}$, que peut-on dire sur les modèles $M_{1/2}$ et M_q ? (utiliser la règle de décision du Tableau 1).
- I.13) Application numérique : Après $n = 200$ lancers, on obtient $r = 115$ fois le côté pile. L'ensemble $\mathcal{Q}$ correspondant est $[0.483, 0.522] \cup [0.627, 0.664]$. Une personne vous assure qu'avec la pièce utilisée, la probabilité d'obtenir "pile" est 1.5 fois supérieure à celle d'obtenir "face". Qu'en pensez-vous?

Connaissant le résultat r de l'expérience, on souhaite s'assurer que la pièce utilisée était bien équilibrée. Pour ce faire, une première méthode consiste à comparer le modèle de la pièce équilibrée au modèle le plus vraisemblable *a posteriori*. Ce modèle (noté M_{opt}) est celui pour lequel $\mathbb{P}([X_i = 1]) = \hat{q}(n, r)$ pour tout $i \in \{1, \dots, n\}$; la valeur $\hat{q}(n, r)$ est la probabilité $q \in [0, 1]$ maximisant la probabilité $\mathbb{P}([S_n = r])$ obtenue sous le modèle $\mathbb{P}([X_i = 1]) = q$ pour tout $i \in \{1, \dots, n\}$.

- I.14) Calculer $\hat{q}(n, 0)$ et $\hat{q}(n, n)$. Donner l'expression de $\hat{q}(n, r)$.
- I.15) Donner l'expression du facteur de Bayes $f_{1/2, opt}(n, r) = \log_{10} p_{1/2}(n, r) - \log_{10} p_{opt}(n, r)$ où $p_{opt}(n, r) := \mathbb{P}([S_n = r])$ lorsque l'on se place sous le modèle M_{opt} .
- I.16) Application numérique : en prenant les mêmes valeurs que dans la question II.5), calculer la valeur du facteur de Bayes $f_{1/2, opt}(n, r)$. Conclure.

Une autre méthode pour s'assurer *a posteriori* de l'équilibre de la pièce consiste à comparer le modèle $M_{1/2}$ au modèle M_{bay} ainsi défini : pour tout $N \in \mathbb{N} \setminus \{0\}$, on définit la variable aléatoire Q_N qui prend ses valeurs dans l'ensemble $\{j/(N+1), j = 1, \dots, N\}$ avec $\mathbb{P}([Q_N = j/(N+1)]) = 1/N$ pour tout $j = 1, \dots, N$. On suppose que la probabilité d'obtenir le résultat r dépend de la valeur de la variable aléatoire Q_N :

$$\mathbb{P}([S_n = r] | [Q_N = j/(N+1)]) = \mathbb{P}\left(\left[\sum_{i=1}^n X_i = r\right]\right),$$

avec $\mathbb{P}([X_i = 1]) = j/(N+1)$ pour tout $i = 1, \dots, n$. On rappelle que la probabilité d'un événement A sachant qu'un événement B de probabilité non nulle est réalisé est donnée par $\mathbb{P}(A | B) = \mathbb{P}(A \cap B) / \mathbb{P}(B)$.

- I.17) En utilisant la formule des probabilités totales, donner l'expression (en fonction de n , r et N) de la probabilité $p_{bay}(n, r)$ que la variable S_n soit égale à r sous le modèle M_{bay} .
- I.18) Application numérique : Pour $n = 200$, $r = 115$ et $N = 100$ on a

$$\log_{10} \left(\frac{1}{N} \sum_{j=1}^N \left(\frac{j}{N+1} \right)^r \left(1 - \frac{j}{N+1} \right)^{n-r} \right) = -60.28$$

Calculer la valeur de facteur de Bayes et conclure.

Problème II – Estimation de la production

Lors de la mise sur le marché de son dernier véhicule, un constructeur automobile a souhaité affecter à chaque voiture vendue un numéro. Ainsi, la première voiture vendue s'est vu attribuer le numéro 1, la deuxième, le numéro 2, etc. Ce numéro est clairement visible à l'arrière du véhicule. Au bout d'un an de mise sur la marché, un concurrent souhaite savoir combien de voitures ont été vendues par ce constructeur automobile. Pour ce faire, il dispose uniquement des numéros des voitures qu'il a croisées lors de l'année écoulée.

Notations : le nombre de voitures observées par le concurrent est noté k et les numéros correspondants, rangés par ordre croissant, sont notés $x_1 < \dots < x_k$. Enfin, le nombre (inconnu) de véhicules vendus est noté N . On a donc forcément $k \leq N$.

On suppose dans la suite que le k -uplet $(x_1 < \dots < x_k)$ est le résultat de l'expérience aléatoire consistant à tirer au hasard et sans remise k numéros parmi $\{1, \dots, n\}$. Cette expérience est modélisée par l'espace $(\Omega_k, \mathcal{A}, \mathbb{P})$ où Ω_k est l'ensemble de tous les k -uplets $(x_1 < \dots < x_k)$ possibles, $\mathcal{A}$ est la tribu de l'ensemble des parties de Ω_k et $\mathbb{P}$ est la probabilité uniforme.

II.1) Quel est (en fonction de N et k) le cardinal de Ω_k .

II.2) Donner l'expression (en fonction de N et k) de la probabilité $\mathbb{P}(\{(x_1 < \dots < x_k)\})$ d'obtenir le tirage $x_1 < \dots < x_k$.

On introduit la variable aléatoire $M_k : \Omega_k \rightarrow \mathbb{N}$ qui à tout k -uplet $(x_1 < \dots < x_k) \in \Omega_k$ associe la plus grande valeur x_k .

II.3) Pour $m < k$ ou $m > N$, quelle est la valeur de $\mathbb{P}([M_k = m])$?

II.4) On suppose que m est tel que $k \leq m \leq N$. Quel est le nombre de k -uplets $(x_1 < \dots < x_k)$ pour lesquels $x_k = m$? En déduire la valeur de $\mathbb{P}([M_k = m])$.

II.5) En utilisant les questions précédentes, montrer que

$$C_N^k := \frac{N!}{k!(N-k)!} = \sum_{m=k}^N C_{m-1}^{k-1}$$

où pour tout entier $\ell > 0$, $\ell! = \ell \times (\ell - 1) \times \dots \times 1$ et avec la convention $0! = 1$.

II.6) Montrer que l'espérance de M_k est donnée par

$$\mathbb{E}(M_k) = \frac{k}{C_N^k} \sum_{m=k}^N C_{m-1}^{k-1}.$$

II.7) En effectuant le changement de variable $\ell = m + 1$ dans la somme ci-dessus et en utilisant le résultat de la question 5), montrer que $\mathbb{E}(M_k) = k(N+1)/(k+1)$.

II.8) On rappelle que le concurrent souhaite estimer le nombre N de véhicules vendus. En utilisant la question précédente, proposer une fonction $\Psi_k(\cdot)$ (indépendante de N évidemment) telle que $\mathbb{E}(\Psi_k(M_k)) = N$. La variable aléatoire $\Psi_k(M_k)$ est un estima-

teur sans biais de N .

II.9) Calculer $\mathbb{E}[(\Psi_k(M_k))^2]$ et en déduire la variance de l'estimateur $\Psi_k(M_k)$ (Aide : remplacer m^2 par $m(m+1) - m$).

II.10) On suppose que $k = N - \lfloor N^a \rfloor$ avec $a \in]0, 1[$, la notation $\lfloor \cdot \rfloor$ étant celle de la partie entière. En utilisant l'inégalité de Bienaymé-Tchebychev (voir le rappel donné à la fin du sujet), montrer que dans ce cas $\Psi_k(M_k) - N$ converge en probabilité vers 0 lorsque $N \rightarrow \infty$.

Problème III – Estimation de la taille d'une population

Pour estimer la taille de la population d'une espèce animale, la méthode de "capture-marquage-recapture" (CMR) est couramment utilisée.

Si on note P une population de $N \in \mathbb{N} \setminus \{0\}$ individus (le nombre N étant inconnu), la méthode CMR consiste à :

- (C) tirer au hasard et sans remise $n \in \{1, \dots, N\}$ individus dans la population P ;
- (M) marquer les n individus et les remettre dans la population P ;
- (R) tirer au hasard et sans remise $m \in \{1, \dots, N\}$ individus dans la population P .

Dans la suite, on raisonnera comme si les N individus de la population étaient numérotés de 1 à N . Ceci n'est évidemment pas le cas en pratique mais permet de simplifier la modélisation du problème sans perte de généralité.

La modélisation probabiliste de la méthode CMR est le triplet $(\Omega, \mathcal{A}, \mathbb{P})$ avec :

- Ω l'ensemble regroupant tous les résultats possible de la méthode CMR. On notera $\omega = (\omega_1, \omega_2)$ un élément de Ω avec
 - ω_1 un vecteur de taille N dont la i -ème composante $\omega_{1,i} = 1$ si l'individu numéro i a été capturé lors de l'étape (C) et 0 sinon;
 - ω_2 un vecteur de taille N dont la i -ème composante $\omega_{2,i} = 1$ si l'individu numéro i a été capturé lors de l'étape (R) et 0 sinon.
- $\mathcal{A}$ la tribu des parties de Ω ;
- $\mathbb{P}$ la probabilité uniforme.

III.1) Donner la valeur des sommes

$$\sum_{i=1}^N \omega_{1,i} \quad \text{et} \quad \sum_{i=1}^N \omega_{2,i}.$$

III.2) Quel est le cardinal de l'ensemble Ω ? En déduire la valeur de la probabilité $\mathbb{P}(\{\omega\})$ pour tout $\omega \in \Omega$.

On introduit la variable aléatoire $K : \Omega \rightarrow \mathbb{N}$ définie par

$$K(\omega) = \sum_{i=1}^N \omega_{1,i} \omega_{2,i}.$$

- III.3) Que peut-on dire lorsque la valeur observée de K est $k \in \mathbb{N}$?
- III.4) Si $n + m > N$, donner les valeurs de k pour lesquelles $\mathbb{P}([K = k]) = 0$.
- III.5) Donner l'ensemble $E(n, m, N) \subset \mathbb{N}$ des valeurs possibles de K .
- III.6) Pour tout $k \in E(n, m, N)$, donner, en fonction de n , m et N l'expression de la probabilité $\mathbb{P}([K = k])$.
- III.7) En déduire la relation
- $$\sum_{k \in E(n, m, N)} C_n^k C_{N-n}^{m-k} = C_N^m.$$
- III.8) En remarquant que $kC_n^k = nC_{n-1}^{k-1}$ et en utilisant la relation ci-dessus, donner l'expression de l'espérance de K .
- III.9) Proposer une fonction $\Psi_{n,m}(\cdot)$ telle que $1/\hat{N} := 1/\Psi_{n,m}(K)$ soit un estimateur sans biais de $1/N$. La variable aléatoire $\hat{N}$ est-elle un estimateur sans biais de N ?

On suppose à présent que $n + m > N$.

- III.10) Donner, en fonction de n , m et N , l'expression de l'espérance de $1/(K+1)$.
- III.11) Utiliser la question précédente pour proposer un estimateur sans biais de N .

Problème IV – Inégalités de Markov et de Cantelli

Dans ce problème, on se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et on note $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire telle que $\mathbb{E}(X) = m \in \mathbb{R}$ et $\mathbb{E}(X^2) = \sigma^2 + m^2$. Pour tout $\varepsilon > 0$, on introduit la variable aléatoire $Y_\varepsilon : \Omega \rightarrow \{0, 1\}$ définie pour tout $\omega \in \Omega$ par

$$Y_\varepsilon(\omega) = \begin{cases} 1 & \text{si } |X(\omega)| > \varepsilon, \\ 0 & \text{si } |X(\omega)| \leq \varepsilon. \end{cases}$$

- IV.1) En posant $p_\varepsilon := \mathbb{P}(|X| > \varepsilon)$, donner l'expression de $\mathbb{E}(Y_\varepsilon)$.
- IV.2) Montrer que pour tout $a \in [0, 2]$, $\mathbb{E}(|X|^a) \geq \mathbb{E}(|X|^a Y_\varepsilon)$.
- IV.3) En déduire l'inégalité de Markov : pour tout $\varepsilon > 0$ et pour tout $a \in [0, 2]$,

$$\mathbb{P}(|X| > \varepsilon) \leq \frac{\mathbb{E}(|X|^a)}{\varepsilon^a}.$$

- IV.4) Montrer que pour tout $\varepsilon > 0$ et $\alpha > 0$, $\mathbb{P}([X - \alpha > \varepsilon]) \leq \mathbb{P}([|X - \alpha| > \varepsilon])$.

IV.5) En utilisant l'inégalité de Markov, montrer que pour tout $\varepsilon > 0$ et $\lambda > 0$,

$$\mathbb{P}([X - m + \lambda > \varepsilon + \lambda]) \leq \frac{\sigma^2 + \lambda^2}{(\varepsilon + \lambda)^2}.$$

IV.6) Après avoir fait l'étude de la fonction $\lambda \in [0, \infty[\mapsto (\sigma^2 + \lambda^2)/(\varepsilon + \lambda)^2$, montrer que pour tout $\varepsilon > 0$, la meilleure majoration est donnée par

$$\mathbb{P}([X - m > \varepsilon]) \leq \frac{\sigma^2}{\varepsilon^2 + \sigma^2}.$$

IV.7) En vous inspirant des questions précédentes, montrer que pour tout $\varepsilon > 0$,

$$\mathbb{P}([X - m < -\varepsilon]) \leq \frac{\sigma^2}{\varepsilon^2 + \sigma^2}.$$

IV.8) En déduire l'inégalité de Cantelli : pour tout $\varepsilon > 0$,

$$\mathbb{P}([|X - m| > \varepsilon]) \leq \frac{2\sigma^2}{\varepsilon^2 + \sigma^2}.$$

IV.9) Pour quelles valeurs de $\varepsilon > 0$, l'inégalité de Cantelli est-elle préférable à celle de Bienaymé-Tchebychev?

Rappel : Pour toute variable aléatoire X de variance σ^2 et d'espérance $m \in \mathbb{R}$, l'inégalité de Bienaymé-Tchebychev est, pour tout $\varepsilon > 0$,

$$\mathbb{P}([|X - m| > \varepsilon]) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2022

Épreuve de mathématiques

Durée : 4h

L'épreuve est constituée de trois exercices indépendants.

Exercice 1. Valeurs prises par une fonction

On note $E = C^0([0, 1], \mathbb{R})$ l'ensemble des fonctions continues de $[0, 1]$ dans $\mathbb{R}$,

$$\mathcal{C} := \left\{ f \in E; \forall t \in [0, 1] f(t) \geq 0 \text{ et } f \text{ est non nulle} \right\}$$

$$\text{et } \mathcal{C}' := \left\{ f \in E; \forall t \in [0, 1] f(t) \geq 0 \text{ et } \int_0^1 f(t) dt = 1 \right\}.$$

On considère, pour tout élément g de E , l'application Φ_g suivante :

$$\Phi_g: \begin{array}{ccc} \mathcal{C} & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \frac{\int_0^1 f(t)g(t) dt}{\int_0^1 f(t) dt}. \end{array}$$

L'objet de l'exercice est de déterminer les valeurs prises par l'application Φ_g .

1. Soit $g \in E$.

(a) Prouver l'égalité : $\Phi_g(\mathcal{C}) = \Phi_g(\mathcal{C}')$.

(b) Soit f_1 et f_2 deux éléments de $\mathcal{C}'$ avec $a_1 = \Phi_g(f_1)$ et $a_2 = \Phi_g(f_2)$.

On suppose $a_1 \leq a_2$. Prouver l'inclusion $[a_1, a_2] \subset \Phi_g(\mathcal{C})$.

2. Dans cette question on suppose que, pour tout réel $t \in [0, 1]$, $g(t) = t$.

(a) Prouver l'inclusion : $\Phi_g(\mathcal{C}) \subset [0, 1]$. Les réels 0 et 1 sont-ils éléments de Φ_g ?

(b) Montrer que $\sup \Phi_g(\mathcal{C}) = 1$.

On pourra, pour tout entier naturel n , calculer $\Phi_g(f_n)$ où $f_n : t \mapsto t^n$.

(c) i. Calculer, pour tout entier naturel n , $\int_0^1 t(1-t)^n dt$.

ii. En déduire que $\inf \Phi_g(\mathcal{C}) = 0$.

(d) Conclure que $\Phi_g(\mathcal{C}) =]0, 1[$.

3. Dans cette question on suppose que, pour tout réel $t \in [0, 1]$, $g(t) = e^t$.

(a) Prouver l'inclusion : $\Phi_g(\mathcal{C}) \subset]1, e[$.

(b) On considère la fonction

$$h: \begin{array}{ccc} \mathbb{R}^* & \longrightarrow & \mathbb{R} \\ \alpha & \longmapsto & \frac{\alpha}{\alpha+1} \frac{e^{\alpha+1} - 1}{e^\alpha - 1}. \end{array}$$

i. Montrer que la fonction h est prolongeable par continuité en 0.

ii. Déterminer $\lim_{\alpha \rightarrow +\infty} h(\alpha)$ et $\lim_{\alpha \rightarrow -\infty} h(\alpha)$.

(c) Prouver l'égalité : $\Phi_g(\mathcal{C}) =]1, e[$.

4. Dans cette question on traite le cas général où l'application g est élément de E et supposée non constante.

- (a) Justifier l'existence d'un couple (a, b) de réels distincts de $[0, 1]$ tel que, pour tout $t \in [0, 1]$, $g(a) \leq g(t) \leq g(b)$.
 On suppose désormais que a et b sont dans $]0, 1[$ et on note N un entier naturel tel que $0 < b - \frac{2}{N} < b + \frac{2}{N} < 1$.

On considère, pour tout entier $n \geq N$, la fonction f_n qui vaut

$$\begin{cases} 0 & \text{si } 0 \leq t \leq b - \frac{1}{n} - \frac{1}{n^2} \text{ ou } b + \frac{1}{n} + \frac{1}{n^2} \leq t \leq 1 \\ 1 & \text{si } b - \frac{1}{n} \leq t \leq b + \frac{1}{n} \end{cases}$$

et qui est affine sur les segments $\left[b - \frac{1}{n} - \frac{1}{n^2}, b - \frac{1}{n}\right]$ et $\left[b + \frac{1}{n}, b + \frac{1}{n} + \frac{1}{n^2}\right]$.

- (b) Représenter le graphe de la fonction f_n .
 (c) Prouver, lorsque l'entier n tend vers $+\infty$, l'équivalence : $\int_0^1 f_n(t) dt \sim \frac{2}{n}$.
 (d) Prouver l'existence d'un réel $K > 0$ (ne dépendant que de g) tel que, pour tout entier $n \geq N$,

$$\int_0^{b-\frac{1}{n}} f_n(t) g(t) dt \leq \frac{K}{n^2} \quad \text{et} \quad \int_{b+\frac{1}{n}}^1 f_n(t) g(t) dt \leq \frac{K}{n^2}.$$

- (e) Soit G une fonction réelle définie au voisinage de b et dérivable en b .

Que vaut $\lim_{h \rightarrow 0} \frac{G(b+h) - G(b-h)}{h}$?

- (f) Prouver l'égalité : $\lim_{n \rightarrow +\infty} \frac{n}{2} \int_{b-\frac{1}{n}}^{b+\frac{1}{n}} g(t) dt = g(b)$.

- (g) En déduire que $\sup \Phi_g(\mathcal{C}) = g(b)$.

On montrerait de même, et on l'admet, que $\inf \Phi_g(\mathcal{C}) = g(a)$.

5. Soit $g \in E$. Que vaut $\Phi_g(\mathcal{C})$?

On cherchera à quelle condition $\max_{t \in [0,1]} g(t)$ est élément de $\Phi_g(\mathcal{C})$.

Exercice 2. Une suite binomiale

Soit n un entier naturel non nul. On note E le $\mathbb{R}$ -espace vectoriel $\mathbb{R}_n[X]$ des polynômes réels de degré au plus n et E^* son dual c'est-à-dire le $\mathbb{R}$ -espace vectoriel des formes linéaires sur E . On note, pour tout entier naturel k , $P^{(k)}$ le polynôme dérivé k -ième d'un polynôme P (avec $P^{(0)} = P$).

On note $P_0 = 1$ et, pour tout entier naturel k non nul, $P_k = X(X-k)^{k-1}$ (on a donc $P_1 = X$).

- Montrer que la famille $\mathcal{B} := (P_0, P_1, \dots, P_n)$ est une base de E .
- On note, pour tout entier r de $\llbracket 0, n \rrbracket$, φ_r l'application qui, à chaque polynôme P de E , associe $\varphi_r(P) = \frac{1}{r!} P^{(r)}(r)$ (on a donc $\varphi_0(P) = P(0)$).

(a) Soit $(k, r) \in \llbracket 0, n \rrbracket$. Calculer $P_k^{(r)}$.

On distinguera les cas $k > r$, $k < r$ et $k = r$.

(b) Calculer, pour tout couple (k, r) d'entiers distincts de $\llbracket 0, n \rrbracket$, la valeur de $\varphi_r(P_k)$.

(c) Calculer, pour tout entier k de $\llbracket 0, n \rrbracket$, la valeur de $\varphi_k(P_k)$.

(d) Soit $P \in E$. Établir l'égalité : $P = \sum_{k=0}^n \varphi_k(P) P_k$.

(e) Montrer que $(\varphi_0, \varphi_1, \dots, \varphi_n)$ est une base de E^* .

3. Établir, pour tout couple (x, y) de nombres réels, l'égalité :

$$(x+y)^n = y^n + x \sum_{k=1}^n \binom{n}{k} (x-k)^{k-1} (y+k)^{n-k}.$$

On pourra considérer le polynôme $(X+y)^n$.

4. Établir, pour tout couple (x, y) de réels, l'égalité :

$$P_n(x+y) = \sum_{k=0}^n \binom{n}{k} P_k(x) P_{n-k}(y).$$

Exercice 3. Étude d'une variable aléatoire

On note, pour tout $(n, p) \in \mathbb{N} \times \mathbb{N}^*$, $a(n, p)$ le nombre de p -listes $(x_1, x_2, \dots, x_p) \in \mathbb{N}^p$ telles que $\sum_{i=1}^p x_i = n$. On a donc $a(n, 1) = 1$.

1. (a) Déterminer, pour tout entier naturel n , la valeur de $a(n, 2)$.

(b) Prouver, pour tout $(n, p) \in \mathbb{N}^2$, l'égalité : $\binom{n+1}{p+1} = \sum_{k=p}^n \binom{n}{k}$.

2. Établir, pour tout $(n, p) \in \mathbb{N} \times \mathbb{N}^*$, l'égalité : $a(n, p+1) = \sum_{k=0}^n a(n-k, p)$.

3. Établir, pour tout $(n, p) \in \mathbb{N} \times \mathbb{N}^*$, l'égalité : $a(n, p) = \binom{n+p-1}{p-1}$.

4. En déduire que, pour tout $(n, p) \in \mathbb{N} \times \mathbb{N}^*$, le nombre de p -listes $(x_1, x_2, \dots, x_p) \in (\mathbb{N}^*)^p$ telles que $\sum_{i=1}^p x_i = n$ vaut $\binom{n-1}{p-1}$.

On munit, pour tout entier naturel n non nul, l'ensemble Ω des parties à n éléments de $\llbracket 1, 2n \rrbracket$ de la tribu pleine et de la probabilité uniforme notée $\mathbf{P}$.

On appelle **bloc** d'une partie A élément de Ω , tout sous-ensemble de A de la forme $\llbracket a, b \rrbracket$ où

$$1 \leq a \leq b \leq 2n \text{ avec } (a=1 \text{ ou } a-1 \notin A) \text{ et } (b=2n \text{ ou } b+1 \notin A).$$

Par exemple, dans le cas où $n=3$, la partie (de $\llbracket 1, 6 \rrbracket$) $\{1, 2, 5\}$ a deux blocs ($\llbracket 1, 2 \rrbracket$ et $\{5\}$) alors que, dans le cas où $n=4$ la partie (de $\llbracket 1, 8 \rrbracket$) $\{2, 4, 6, 7\}$ a trois blocs ($\{2\}$, $\{4\}$ et $\llbracket 6, 7 \rrbracket$).

On note, pour tout entier naturel n non nul, X_n la variable aléatoire qui, à chaque partie A à n éléments de $\llbracket 1, 2n \rrbracket$, associe le nombre de blocs de A . La variable X_1 est donc constante égale à 1.

5. (a) Déterminer la loi de X_2 et son espérance.

(b) Déterminer la loi de X_3 et son espérance.

6. Soit $n \in \mathbb{N}^*$.

(a) Calculer $\mathbf{P}(X_n = 1)$ et $\mathbf{P}(X_n = n)$.

(b) Soit $k \in \llbracket 1, n \rrbracket$. Montrer que la donnée d'une partie A à n éléments de $\llbracket 1, 2n \rrbracket$ avec k blocs équivaut à la donnée d'une liste $(j_0, i_1, j_1, i_2, j_2, \dots, i_{k-1}, j_{k-1}, i_k, j_k)$ où $i_1, i_2, \dots, i_k, j_1, j_2, \dots, j_{k-1}$ sont des entiers naturels non nuls et j_0 et j_k des entiers naturels vérifiant les conditions

$$\sum_{r=1}^k i_r = n \quad \text{et} \quad \sum_{r=0}^k j_r = n.$$

(c) En déduire, pour tout entier $k \in \llbracket 1, n \rrbracket$, l'égalité : $\mathbf{P}(X_n = k) = \frac{\binom{n-1}{k-1} \binom{n+1}{k}}{\binom{2n}{n}}$.

(d) i. Calculer l'espérance de la variable X_n .

ii. Calculer la variance de la variable X_n .

7. Prouver, pour tout $\varepsilon > 0$, l'égalité :

$$\lim_{n \rightarrow +\infty} \mathbf{P}\left(\left|\frac{X_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right) = 0.$$

8. Soit $n \in \mathbb{N}^*$. On lance $2n$ fois une pièce équilibrée et on note $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{2n})$ la suite de résultats obtenus (avec $\varepsilon_i = 1$ (resp. 0) si le i -ème lancer a donné pile (resp. face)). Quelle est la probabilité d'obtenir une telle suite comportant n fois le 1 avec exactement k blocs de 1?

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2022

Épreuve à option (A) : Mathématiques

Durée : 4h

On rappelle que la **fonction Gamma** d'Euler est définie par

$$\forall x > 0, \quad \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt \quad (1)$$

qu'elle est continue sur $\mathbb{R}_+^*$ et qu'elle vérifie la formule récursive $\Gamma(x+1) = x\Gamma(x)$, valable pour tout réel strictement positif x .

Cette fonction intervient à maintes reprises dans la suite.

L'énoncé est divisé en quatre parties largement indépendantes, que les candidats ne sont pas tenus de traiter dans l'ordre.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

Partie I Un lien avec loi de Poisson

Dans cette partie, on considère une suite $(X_n)_{n \in \mathbb{N}^*}$ variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et on suppose que, pour tout $n \in \mathbb{N}^*$, X_n suit la loi de Poisson de paramètre n .

1. Justifier que, pour tout $n \in \mathbb{N}^*$, la probabilité $\mathbf{P}([X_n > k])$ tend vers 0 quand l'entier k tend vers l'infini.

2. Soit $n \in \mathbb{N}^*$.

- a) Justifier, pour tout entier naturel k , l'égalité :

$$\mathbf{P}([X_n > k]) = \mathbf{P}([X_n > k-1]) - \frac{n^k}{k!} e^{-n}.$$

- b) En utilisant le résultat précédent et une intégration par parties, établir que :

$$\forall k \in \mathbb{N}, \quad \mathbf{P}([X_n > k]) = \frac{1}{k!} \int_0^n t^k e^{-t} dt \quad (2)$$

3. Soit $k \in \mathbb{N}$.

- a) Préciser la limite de $\mathbf{P}([X_n > k])$ quand n tend vers l'infini.

- b) En utilisant (2), justifier que $\mathbf{P}([X_n \leq k])$ est équivalent à $\frac{n^k}{k!} e^{-n}$ quand n tend vers l'infini.

4. Pour tout $n \in \mathbb{N}^*$, on pose : $I_n = \int_0^n t^{2n} e^{-t} dt$.

- a) Justifier l'encadrement : $\frac{n^{2n+1}}{2n+1} e^{-n} \leq I_n \leq \frac{n^{2n+1}}{2n+1}$.

- b) En déduire la limite de I_n quand n tend vers l'infini.

- c) Démontrer que I_n est négligeable devant $(2n)!$ quand n tend vers l'infini (on pourra utiliser la propriété (2)).

Partie II La fonction Γ comme transformée intégrale

On note S l'ensemble des fonctions réelles f définies et de classe C^∞ sur $\mathbb{R}_+$, telles que, pour tout couple (n, p) de nombres entiers positifs ou nuls, la fonction $t \mapsto t^p f^{(n)}(t)$ est bornée sur $\mathbb{R}_+$.

5. Pour quelles valeurs réelles de a la fonction $t \mapsto e^{at}$ est-elle un élément de S ?

6. Soit k un nombre entier strictement positif et f_k la fonction définie sur $\mathbb{R}_+$ par :

$$\forall t \geq 0, \quad f_k(t) = \exp(-t^k).$$

a) Montrer que, pour tout entier positif n , il existe un polynôme $P_{k,n}$ de $\mathbb{R}[X]$ tel que la dérivée n -ième de f_k vérifie :

$$\forall t \geq 0, \quad f_k^{(n)}(t) = P_{k,n}(t) f_k(t).$$

b) En déduire que f_k est un élément de S .

7. Montrer que, pour tout élément f de S et tout réel $x > 0$, la fonction $t \mapsto t^{x-1} f(t)$ est intégrable sur $]0, +\infty[$.

Pour tout $f \in S$, on définit sur $\mathbb{R}_+^*$ la fonction $G(f)$ par :

$$\forall x > 0, \quad G(f)(x) = \int_0^{+\infty} t^{x-1} f(t) dt \quad (3)$$

8. a) Exprimer, pour tout entier strictement positif k , la fonction $G(f_k)$ à l'aide de la fonction Gamma.

b) Donner un équivalent de la fonction Γ en 0 et en déduire que, pour tout $x > 0$, $G(f_k)(x)$ tend vers $\frac{1}{x}$ lorsque k tend vers l'infini.

9. Montrer que, pour tout élément f de S , la fonction $G(f)$ est de classe C^∞ sur $]0, +\infty[$.

10. Dans cette question, on suppose que f est une fonction à valeurs strictement positives qui appartient à S .

a) Montrer que $(G(f))(x)$ tend vers $+\infty$ quand x tend vers $+\infty$.

b) Montrer que $(G(f))(x)$ est équivalent à $f(0)/x$ quand x tend vers 0.

c) Montrer que la fonction $G(f)$ admet un minimum global, atteint en un point unique.

Partie III Application au prolongement de la fonction ζ de Riemann

Pour tout élément f de l'espace S défini dans la deuxième partie du problème, on associe la fonction $Z(f)$ définie sur $]0, +\infty[$ par :

$$\forall x > 0, \quad Z(f)(x) = \frac{1}{\Gamma(x)} \int_0^{+\infty} t^{x-1} f(t) dt \quad (4)$$

11. Soit f un élément de S .

a) Montrer que la dérivée f' de f appartient à S et vérifie :

$$\forall x > 0, \quad G(f')(x+1) = -x G(f)(x).$$

b) En déduire que, pour tout $n \in \mathbb{N}$, on a :

$$\forall x > 0, \quad Z(f^{(n)})(x+n) = (-1)^n Z(f)(x).$$

c) Montrer qu'on peut définir, de manière cohérente, un prolongement $\overline{Z(f)}$ de $Z(f)$ à $\mathbb{R}$, en posant, pour tout nombre réel x et tout nombre entier n tel que $n > -x$:

$$\overline{Z(f)}(x) = (-1)^n Z(f^{(n)})(x+n).$$

d) Montrer que $\overline{Z(f)}$, ainsi défini, est une fonction de classe C^∞ sur $\mathbb{R}$ et vérifie, pour tout $n \in \mathbb{N}$:

$$\overline{Z(f)}(-n) = (-1)^n f^{(n)}(0).$$

12. Dans cette question, on note f la fonction définie sur $\mathbb{R}_+$ par : $f(t) = \begin{cases} \frac{t}{e^t - 1} & \text{si } t > 0 \\ 1 & \text{si } t = 0 \end{cases}$.

a) Justifier la validité des deux développements en série suivants :

$$\begin{aligned} \text{(i)} \quad \forall t > 0, \quad f(t) &= t e^{-t} \sum_{n=0}^{+\infty} e^{-nt} \\ \text{(ii)} \quad \forall t \geq 0, \quad f(t) &= \frac{1}{1 + \sum_{n=1}^{+\infty} \frac{t^n}{(n+1)!}}. \end{aligned}$$

b) Montrer que f est de classe C^∞ sur $\mathbb{R}_+$.

c) Pour tout nombre entier strictement positif k , montrer que la série de terme général $n^k e^{-n}$ est convergente et que sa somme $\sum_{n=1}^{+\infty} n^k e^{-n}$ majore la valeur absolue de la dérivée k -ième de la fonction $t \mapsto \sum_{n=0}^{+\infty} e^{-nt}$ sur l'intervalle fermé $[1, +\infty[$.

d) En déduire que f est un élément de S .

13. On rappelle que la **fonction Zêta** de Riemann est définie par :

$$\forall x > 1, \quad \zeta(x) = \sum_{n=1}^{+\infty} \frac{1}{n^x} \quad (5)$$

a) Démontrer que, pour l'élément f de S défini dans la question précédente, on a :

$$\forall x > 0, \quad Z(f)(x) = x \zeta(x+1).$$

b) La propriété précédente permet de prolonger la fonction ζ à la droite réelle privée de 1 en posant :

$$\forall x \neq 1, \quad \bar{\zeta}(x) = \frac{1}{x-1} \overline{Z(f)}(x-1).$$

En utilisant un développement limité de f , calculer ses valeurs en $\overline{Z(f)}(0)$ et $\overline{Z(f)}(-1)$.

Partie IV Un autre prolongement

Dans cette partie, P et Q désignent deux polynômes de $\mathbb{R}[X]$, non nuls et **premiers entre eux**, et on note $E(P, Q)$ l'ensemble des solutions sur $\mathbb{R}$ de l'équation différentielle :

$$P y'' + Q y' + (Q - P) y = 0 \quad (6)$$

Autrement dit, $E(P, Q)$ est l'ensemble des fonctions réelles f deux fois dérivables sur $\mathbb{R}$ telles que :

$$\forall x \in \mathbb{R}, \quad P(x) f''(x) + Q(x) f'(x) + (Q(x) - P(x)) f(x) = 0.$$

14. a) Montrer que $E(P, Q)$ est un $\mathbb{R}$ -espace vectoriel qui contient la fonction $x \mapsto e^{-x}$.

b) Montrer que, si $E(P, Q)$ contient une fonction de la forme $x \mapsto e^{\lambda x}$ avec $\lambda \neq -1$, les polynômes P et Q sont nécessairement constants.

c) Trouver $E(P, Q)$ lorsque P et Q sont constants (non nuls).

15. On suppose, dans cette question, que le polynôme P possède une unique racine réelle a et on note Φ l'application linéaire de $E(P, Q)$ dans $\mathbb{R}^4$ qui associe à tout élément f de $E(P, Q)$ le vecteur :

$$\Phi(f) = (f(a-1), f'(a-1), f(a+1), f'(a+1)).$$

a) En utilisant le théorème de Cauchy linéaire, montrer que l'application Φ est injective.

b) Montrer qu'il n'existe pas d'élément f de $E(P, Q)$ vérifiant :

$$\begin{cases} f(a-1) = f'(a-1) = 0 \\ f(a+1) = -f'(a+1) = e^{-(a+1)} \end{cases}.$$

c) Dédire des résultats précédents que la dimension de l'espace vectoriel $E(P, Q)$ est au plus égale à 3.

d) Montrer que, si $P(X) = X^3$ et $Q(X) = -1$, la dimension de l'espace vectoriel $E(P, Q)$ est égale à 3.

16. a) Montrer que la fonction $\Delta : x \mapsto \frac{\Gamma(\frac{x}{2})}{\Gamma(x)}$, définie sur $]0, +\infty[$, vérifie la formule récurrente :

$$\forall x > 0, \quad \Delta(x+2) = \frac{\Delta(x)}{2(x+1)}.$$

b) En utilisant la fonction $f_2 : t \mapsto \exp(-t^2)$ définie dans la deuxième partie, montrer que la fonction Δ admet un prolongement $\bar{\Delta}$ de classe C^∞ sur $\mathbb{R}$, vérifiant :

$$\forall x \in \mathbb{R}, \quad \bar{\Delta}(x) = 2(x+1)\bar{\Delta}(x+2).$$

c) Comment utiliser une équation différentielle de la forme (6) pour trouver une formule récurrente satisfaite à la fois par les fonctions constantes et par la fonction $\bar{\Delta}$?

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2022

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet est composé d'un problème, divisé en trois parties, et d'un exercice qui peuvent être traités indépendamment. La notation tiendra largement compte de la clarté et de la précision des réponses.

Quelques rappels

- Un espace probabilisé est un triplet $(\Omega, \mathcal{A}, \mathbb{P})$ où Ω est un ensemble quelconque appelé l'univers, $\mathcal{A}$ est une tribu de partie de Ω et $\mathbb{P}$ est une mesure de probabilité sur l'espace $(\Omega, \mathcal{A})$. Les éléments de $\mathcal{A}$ sont des sous-ensembles de Ω appelés événements.
- Dans le cas où l'univers Ω est de cardinal fini, on prend $\mathcal{A} = \mathcal{P}(\Omega)$, l'ensemble contenant tous les sous-ensembles de Ω . Il est à noter que l'ensemble vide $\emptyset$ est un élément de $\mathcal{P}(\Omega)$. La mesure de probabilité $\mathbb{P}$ est alors entièrement caractérisée par l'ensemble des valeurs $\{\mathbb{P}(\{\omega\}) \mid \omega \in \Omega\}$ et on a

$$\sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 1.$$

- Une variable aléatoire discrète X est une fonction de Ω dans un ensemble fini ou dénombrable E et telle que pour tout $x \in E$, $\{\omega \in \Omega \mid X(\omega) = x\} \in \mathcal{A}$. Son espérance est donnée par

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) X(\omega).$$

- **Inégalité de Markov** – Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Pour tout $\varepsilon > 0$ et $k \in \mathbb{N}$,

$$\mathbb{P}(|X| \geq \varepsilon) \leq \frac{\mathbb{E}(|X|^k)}{\varepsilon^k}.$$

Problème – Estimation d'une moyenne par plan de sondage

Dans ce problème, on note $U = \{u_1, \dots, u_N\}$ une population de $N \in \mathbb{N} \setminus \{0\}$ individus. Par exemple, U peut être l'ensemble des $N = 16,7$ millions retraités en France (pour l'année 2021). A chaque individu u_i , $i \in \{1, \dots, N\}$, est associée une valeur numérique y_i (par exemple le revenu annuel perçut en 2021 par le retraité u_i). La moyenne des valeurs $\{y_i \mid i = 1, \dots, N\}$ est donnée par

$$\bar{y} := \frac{1}{N} \sum_{i=1}^N y_i.$$

Cette valeur n'est la plupart du temps pas accessible en pratique. Dans l'exemple du revenu annuel des retraités, il est difficile voire impossible de croiser les différentes sources de revenu d'un individu donné et donc de connaître l'ensemble des valeurs $\{y_i \mid i = 1, \dots, N\}$. On doit donc se contenter d'estimer la moyenne en effectuant un sondage auprès de la population. Plus précisément, on choisit aléatoirement un échantillon (ou base de sondage) $S \subset U$ et on estime le revenu moyen $\bar{y}$ sur la base des valeurs y_i fournies par les individus appartenant à la base de sondage S .

De manière plus formelle, l'échantillon S est choisi selon un plan de sondage qui est un espace probabilisé $(\mathcal{P}(U), \mathcal{A}, \mathbb{P})$ où l'univers $\mathcal{P}(U)$ est l'ensemble des parties de U (c'est-à-dire l'ensemble de tous les sous-ensembles de U), $\mathcal{A}$ est l'ensemble des parties de $\mathcal{P}(U)$ et enfin $\mathbb{P}$ est une mesure de probabilité sur $(\mathcal{P}(U), \mathcal{A})$. La probabilité de choisir un échantillon $S \subset U$ donné est donc $\mathbb{P}(\{S\})$.

Important – Il faut bien comprendre que dans un sondage, les valeurs y_i , $i = 1, \dots, N$ ne sont pas aléatoires. L'aléa provient uniquement de choix de l'échantillon par le plan de sondage.

Partie I : Quelques plans de sondage

Exemple 1 On suppose pour commencer que $U = \{u_1, u_2, u_3\}$.

- I.1) Donner les 8 éléments de l'ensemble $\mathcal{P}(U)$.
- I.2) Donner un élément de $\mathcal{A}$ (celui que vous voulez)
- I.3) Quelle est la valeur de la somme

$$\sum_{S \in \mathcal{P}(U)} \mathbb{P}(\{S\}).$$

On suppose que pour tout $S \in \mathcal{P}(U)$, $\mathbb{P}(\{S\}) = \text{card}(S)/c_3$ où c_3 est une constante strictement positive.

- I.4) Donner la valeur de c_3 .

On se place à présent dans le cas plus général où $U = \{u_1, \dots, u_N\}$ avec $N \in \mathbb{N} \setminus \{0\}$.

- I.5) Quel est, en fonction de N , le cardinal de $\mathcal{P}(U)$?

On suppose que pour tout $S \in \mathcal{P}(U)$, $\mathbb{P}(\{S\}) = \text{card}(S)/c_N$.

- I.6) Donner, en fonction de N , la valeur de c_N où c_N est une constante strictement positive (vous donnerez l'expression la plus simple possible de cette constante).

Exemple 2 – Plan de sondage aléatoire simple (plan SAS) Soit $n \in \{1, \dots, N\}$. Le plan SAS est l'espace probabilisé $(\mathcal{P}(U), \mathcal{A}, \mathbb{P}_n^{(SAS)})$ où la probabilité $\mathbb{P}_n^{(SAS)}$ est définie pour tout échantillon $S \subset U$ par

$$\mathbb{P}_n^{(SAS)}(\{S\}) := \begin{cases} 0 & \text{si } \text{card}(S) \neq n, \\ p_n & \text{si } \text{card}(S) = n. \end{cases}$$

- I.7) Donner la valeur de p_1 (pour le cas $n = 1$) ainsi que la valeur de p_N (pour le cas $n = N$) qu'il faut prendre pour que $\mathbb{P}_n^{(SAS)}$ soit bien une mesure de probabilité.
- I.8) Donner, en fonction de N et n , la valeur de p_n .

Une méthode pour choisir aléatoirement un échantillon $S \subset U$ selon le plan SAS est décrite dans l'algorithme suivant :

Etape 1 – On choisit au hasard un individu dans la population U , chaque individu ayant la même probabilité d'être choisi. Cet individu est noté u_1^ .*

Etape 2 – On choisit au hasard un individu dans la population $U \setminus \{u_1^\}$, chaque individu ayant la même probabilité d'être choisi. Cet individu est noté u_2^* .*

...

Etape n – On choisit au hasard un individu dans la population $U \setminus \{u_1^, \dots, u_{n-1}^*\}$, chaque individu ayant la même probabilité d'être choisi. Cet individu est noté u_n^* .*

A l'issue de l'étape n , on obtient donc l'échantillon $\{u_1^*, \dots, u_n^*\}$ de taille n .

I.9) Calculer la probabilité de choisir l'individu u_1 lors de la première étape.

I.10) Calculer la probabilité de choisir l'individu u_1 lors de la première étape et l'individu u_2 lors de la deuxième étape.

I.11) Calculer la probabilité que les individus u_1 et u_2 soient choisis à l'issue de la deuxième étape.

I.12) En déduire que la probabilité que l'échantillon $S = \{u_1, \dots, u_n\}$ soit choisi par cet algorithme est égale à p_n .

Exemple 3 – Plan de sondage de Bernoulli L'algorithme suivant permet de choisir aléatoirement un échantillon selon le plan de sondage de Bernoulli.

Etape 1 – On note $x_1, \dots, x_N$ les réalisations de N variables aléatoires $X_1, \dots, X_N$ indépendantes et de même loi de Bernoulli de paramètre $p \in]0, 1[$.

Etape 2 – On construit l'échantillon $S \subset U$ en prenant les individus u_i pour lesquels $x_i = 1$. La taille de l'échantillon ainsi obtenu est évidemment aléatoire.

I.13) Quelle est la probabilité d'obtenir l'ensemble vide à l'issue de cet algorithme?

I.14) Pour tout $k \in \{0, \dots, N\}$, donner l'expression de la probabilité d'obtenir un échantillon de taille k en fonction de N , p et k (justifier votre réponse).

I.15) Pour $k \in \{0, \dots, N\}$, quelle est la probabilité d'obtenir l'échantillon $\{u_1, \dots, u_k\}$?

Le plan de sondage de Bernoulli est donc l'espace probabilisé $(\mathcal{P}(U), \mathcal{A}, \mathbb{P}_p^{(Ber)})$ où la probabilité $\mathbb{P}_p^{(Ber)}$ est définie pour tout $S \in \mathcal{P}(U)$ par

$$\mathbb{P}_p^{(Ber)}(\{S\}) = p^{\text{card}(S)}(1-p)^{N-\text{card}(S)}.$$

On introduit la variable aléatoire Z_p définie sur l'espace probabilisé $(\mathcal{P}(U), \mathcal{A}, \mathbb{P}_p^{(Ber)})$ et telle que pour tout $S \in \mathcal{P}(U)$, $Z_p(S) = \text{card}(S)$.

I.16) Calculer l'espérance de Z_p (vous donnerez le résultat sous sa forme la plus simple possible).

Exemple 4 – Plan de sondage systématique Pour ce plan de sondage, l'échantillon est choisi aléatoirement de la façon suivante : pour un pas d'échantillonnage $a \in \{1, \dots, N\}$ fixé par l'utilisateur, on choisit aléatoirement avec équiprobabilité un entier r dans l'ensemble $\{1, \dots, a\}$. On obtient alors l'échantillon

$$\{u_r, u_{r+a}, \dots, u_{r+(n-1)a}\} = \{u_{r+ja} \mid j \in \{0, \dots, n-1\}\},$$

où $n = \lfloor (N - r) / a \rfloor + 1$, la fonction $\lfloor \cdot \rfloor$ étant la partie entière inférieure.

Afin de se familiariser avec ce plan de sondage, on suppose dans un premier temps que $U = \{u_1, \dots, u_{10}\}$ et que le pas d'échantillonnage est $a = 3$.

I.17) Donner l'ensemble des échantillons que l'on peut obtenir avec le plan de sondage systématique.

I.18) Quelle est la probabilité d'obtenir l'échantillon $\{u_1, u_5, u_8\}$? Celle d'obtenir l'échantillon $\{u_2, u_5, u_8\}$?

On se place à présent dans le cadre général d'une population U de taille N et d'un plan systématique de pas d'échantillonnage $a \in \{1, \dots, N\}$.

I.19) Quelle est la probabilité $\mathbb{P}_a^{(Sys)}(\{S\})$ d'obtenir l'échantillon $S \subset U$.

Partie II – Probabilité d'inclusion

On s'intéresse ici au calcul de la probabilité d'inclusion d'un ou plusieurs individus pour un plan de sondage $(\mathcal{P}(U), \mathcal{A}, \mathbb{P})$.

Pour un individu $u_k \in U$ donné, sa probabilité d'inclusion est la probabilité $\pi^{(k)}$ qu'un échantillon S tiré selon le plan de sondage contienne u_k . Autrement dit,

$$\pi^{(k)} := \mathbb{P}(\{S \in \mathcal{P}(U) \mid u_k \in S\}).$$

On note $A_k \in \mathcal{A}$ l'événement $\{S \in \mathcal{P}(U) \mid u_k \in S\}$. On rappelle que $\pi^{(k)} = \mathbb{P}(A_k) = \mathbb{E}(\mathbb{I}_{A_k})$ où $\mathbb{I}_{A_k}$ est la variable aléatoire définie sur $(\mathcal{P}(U), \mathcal{A}, \mathbb{P})$ et telle que pour tout $S \in \mathcal{P}(U)$,

$$\mathbb{I}_{A_k}(S) = \begin{cases} 1 & \text{si } S \in A_k, \\ 0 & \text{si } S \notin A_k. \end{cases}$$

Pour tout couple $(k, \ell) \in \{1, \dots, N\}^2$ avec $k \neq \ell$, la probabilité d'inclusion des individus u_k et u_ℓ est donnée par

$$\pi^{(k, \ell)} := \mathbb{P}(\{S \in \mathcal{P}(U) \mid \{u_k, u_\ell\} \subset S\}) = \mathbb{P}(A_k \cap A_\ell) = \mathbb{E}(\mathbb{I}_{A_k \cap A_\ell}).$$

II.1) En utilisant la formule donnant l'espérance d'une variable aléatoire (voir les rappels en début de sujet) donner, pour tout $(k, \ell) \in \{1, \dots, N\}^2$ avec $k \neq \ell$, les expressions des probabilités $\pi^{(k)}$ et $\pi^{(k, \ell)}$.

II.2) Pour le plan de sondage SAS (exemple 2), donner les expressions de $\pi^{(k)}$ et $\pi^{(k, \ell)}$ en fonction de n et N .

II.3) Pour le plan de sondage de Bernoulli (exemple 3), montrer en utilisant la réponse à la question II.1) que pour tout $k \in \{1, \dots, N\}$,

$$\pi^{(k)} = \sum_{n=1}^N p^n (1-p)^{N-n} \sum_{S \in B_n} \mathbb{I}_{A_k}(S),$$

où $B_n = \{S \in \mathcal{P}(U) \mid \text{card}(S) = n\} \in \mathcal{A}$.

II.4) Donner, en fonction de n et N , l'expression de la somme

$$\sum_{S \in B_n} \mathbb{I}_{A_k}(S),$$

et en déduire l'expression de $\pi^{(k)}$ pour le plan de sondage de Bernoulli.

II.5) En vous inspirant des questions II.3) et II.4), donner l'expression de $\pi^{(k,\ell)}$ pour le plan de sondage de Bernoulli.

II.6) Pour le plan de sondage systématique (exemple 4), donner l'expression de $\pi^{(k)}$ en fonction du pas d'échantillonnage a .

Partie III – Estimation de la moyenne

L'estimation de la moyenne $\bar{y}$ définie dans l'introduction du problème passe par l'estimation de la somme

$$t_N = \sum_{i=1}^N y_i.$$

Pour estimer t_N , on dispose uniquement des valeurs y_i pour les individus u_i qui appartiennent à l'échantillon choisi selon le plan de sondage $(\mathcal{P}(U), \mathcal{A}, \mathbb{P})$. On introduit dans un premier temps la variable aléatoire $\tilde{t}_N$ définie sur l'espace probabilisé $(\mathcal{P}(U), \mathcal{A}, \mathbb{P})$ telle que pour tout $S \in \mathcal{P}(U)$,

$$\tilde{t}_N(S) = \sum_{u_k \in S} y_k = \sum_{k=1}^N y_k \mathbb{I}_{A_k}(S),$$

où $A_k = \{S \in \mathcal{P}(U) \mid u_k \in S\}$. On suppose dans toute la suite que $\pi^{(k)} = \mathbb{P}(A_k) > 0$ pour tout $k \in \{1, \dots, N\}$.

III.1) Montrer que

$$\mathbb{E}(\tilde{t}_N) = \sum_{k=1}^N \pi^{(k)} y_k$$

III.2) En déduire un estimateur sans biais de la moyenne $\bar{y}$.

III.3) Donner l'expression de la variance de l'estimateur sans biais trouvé à la question précédente.

Exercice – Inégalité de Hoeffding

Soient $X_1, \dots, X_n$ des variables aléatoires réelles définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose pour cet exercice que

- Les variables aléatoires $X_1, \dots, X_n$ sont indépendantes;
- Pour tout $i \in \{1, \dots, n\}$, $\mathbb{E}(X_i) = 0$.

- Il existe $-\infty < a < 0 < b < +\infty$ tels que

$$a \leq \inf_{\omega \in \Omega} X(\omega) < \sup_{\omega \in \Omega} X(\omega) \leq b.$$

En posant $S_n = X_1 + \dots + X_n$, l'objectif de cet exercice est d'obtenir l'inégalité de Hoeffding donnée pour tout $\varepsilon > 0$ par

$$\mathbb{P}(|S_n| \geq \varepsilon) \leq 2 \exp\left(-\frac{2\varepsilon^2}{n(b-a)^2}\right).$$

- a) En utilisant la convexité de la fonction exponentielle, montrer que pour tout $t \in \mathbb{R}$ et $x \in [a, b]$,

$$\exp(tx) \leq \frac{b-x}{b-a} \exp(ta) + \frac{x-a}{b-a} \exp(tb)$$

- b) En déduire que pour tout $i \in \{1, \dots, n\}$ et pour tout $t \in \mathbb{R}$,

$$\mathbb{E}(\exp(tX_i)) \leq \exp\left(g((b-a)t)\right),$$

où $g: \mathbb{R} \rightarrow \mathbb{R}$ est la fonction définie pour tout $y \in \mathbb{R}$ par

$$g(y) = \frac{a}{b-a}y + \log\left(\frac{b - a \exp(y)}{b-a}\right)$$

- c) En utilisant un développement limité, montrer que pour tout $y \in \mathbb{R}$,

$$g(y) \leq \frac{y^2}{2} \sup_{y \in \mathbb{R}} g''(y),$$

où g'' est la dérivée seconde de g .

On admettra pour la suite que

$$\sup_{y \in \mathbb{R}} g''(y) \leq \frac{1}{4}.$$

- d) Déduire des questions précédentes que pour tout $t \in \mathbb{R}$,

$$\mathbb{E}(\exp(tS_n)) \leq \exp\left(\frac{nt^2(b-a)^2}{8}\right).$$

- e) En utilisant l'inégalité de Markov (voir les rappels au début du sujet), montrer que pour tout $t > 0$ et $\varepsilon > 0$,

$$\mathbb{P}(S_n \geq \varepsilon) \leq \frac{\mathbb{E}(\exp(tS_n))}{\exp(t\varepsilon)}.$$

- f) En prenant $t = d\varepsilon/n$ avec $d > 0$, trouver la constante $M_n(a, b, d, \varepsilon)$ telle que

$$\mathbb{P}(S_n \geq \varepsilon) \leq M_n(a, b, d, \varepsilon).$$

g) Quelle valeur de $d > 0$ (qui dépendra de a et b) faut-il prendre pour obtenir l'inégalité

$$\mathbb{P}([S_n \geq \varepsilon]) \leq \exp\left(-\frac{2\varepsilon^2}{n(b-a)^2}\right).$$

h) Expliquer pourquoi on a également l'inégalité

$$\mathbb{P}([S_n \leq -\varepsilon]) \leq \exp\left(-\frac{2\varepsilon^2}{n(b-a)^2}\right).$$

i) Dédurre l'inégalité de Hoeffding des questions précédentes.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2023

Épreuve de mathématiques

Durée : 4h

L'épreuve est constituée de deux exercices indépendants.

Exercice 1. Une inégalité entre sommes de séries

On note $\mathcal{E}$ l'ensemble des suites réelles (indexées par $\mathbb{N}^*$) à termes strictement positifs telles que la série $\sum a_n$ converge. On pose, pour tout élément $(a_n)_{n \in \mathbb{N}^*}$ de $\mathcal{E}$ et pour tout entier n non nul,

$$h_n = \frac{n}{\sum_{k=1}^n \frac{1}{a_k}}.$$

L'objet de l'exercice est de prouver la convergence de la série $\sum h_n$ et de comparer sa somme à celle de la série $\sum a_n$.

1. Un premier exemple

On pose, dans cette question, pour tout entier naturel n non nul, $a_n = \frac{1}{n(n+1)}$.

- (a) Montrer que la série $\sum a_n$ converge et déterminer la somme $\sum_{n=1}^{+\infty} a_n$.
- (b)
 - i. Calculer, pour tout entier naturel n non nul, la valeur de h_n .
 - ii. Établir la convergence de la série $\sum h_n$ et déterminer la somme $\sum_{n=1}^{+\infty} h_n$.

2. Un second exemple

Soit q un réel de $]0, 1[$. On pose, pour tout entier naturel n non nul, $a_n = q^{n-1}$.

- (a) Indiquer la valeur de la somme $\sum_{n=1}^{+\infty} a_n$ et pour tout entier naturel n non nul, la valeur de h_n .
- (b) Établir la convergence de la série $\sum h_n$ et prouver la majoration : $\sum_{n=1}^{+\infty} h_n \leq \frac{1}{(1-q)^2}$.

3. Soit n un entier non nul et soit $(x_1, x_2, \dots, x_n)$ et $(y_1, y_2, \dots, y_n)$ deux n -uplets de nombres réels.

(a) Prouver l'égalité : $\left(\sum_{i=1}^n x_i y_i\right)^2 + \sum_{1 \leq i < j \leq n} (x_i y_j - x_j y_i)^2 = \left(\sum_{i=1}^n x_i^2\right) \left(\sum_{i=1}^n y_i^2\right).$

(b) En déduire l'inégalité de Cauchy-Schwarz : $\left|\sum_{i=1}^n x_i y_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}.$

4. Prouver, pour tout entier naturel k non nul, l'inégalité : $\sum_{n=k}^{+\infty} \frac{1}{n(n+1)^2} \leq \frac{1}{2k^2}.$

On s'intéressera à la monotonie de la suite de terme général $u_k = \frac{1}{2k^2} - \sum_{n=k}^{+\infty} \frac{1}{n(n+1)^2}.$

5. Soit $(a_n)_{n \in \mathbb{N}^*}$ un élément de $\mathcal{E}$.

- (a) Prouver, pour tout entier naturel n non nul, l'inégalité :

$$\left(\frac{n(n+1)}{2}\right)^2 \leq \left(\sum_{k=1}^n k^2 a_k\right) \left(\sum_{k=1}^n \frac{1}{a_k}\right).$$

- (b) En déduire, pour tout entier naturel p non nul, l'inégalité :

$$\sum_{n=1}^p h_n \leq 4 \sum_{n=1}^p \frac{1}{n(n+1)^2} \left(\sum_{k=1}^n k^2 a_k\right).$$

- (c) Prouver, pour tout entier naturel p non nul, l'inégalité : $\sum_{n=1}^p h_n \leq 2 \sum_{k=1}^p a_k.$

- (d) En déduire la convergence de la série $\sum h_n$ et l'inégalité : $\sum_{n=1}^{+\infty} h_n \leq 2 \sum_{n=1}^{+\infty} a_n.$

6. Soit C un réel strictement positif tel que, pour tout élément $(a_n)_{n \in \mathbb{N}^*}$ de $\mathcal{E}$, on a

$$\sum_{n=1}^{+\infty} h_n \leq C \sum_{n=1}^{+\infty} a_n.$$

On va montrer que C est au moins égal à 2.

On pose, pour tout entier naturel n non nul, $a_n = \frac{1}{n^\alpha}$ où α est un réel strictement supérieur à 1 et on rappelle qu'on dispose de l'égalité : $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$

(a) Prouver l'inégalité : $\sum_{n=1}^{+\infty} \frac{n}{(n+1)^{\alpha+1}} \geq \sum_{n=1}^{+\infty} \frac{1}{n^\alpha} - \frac{\pi^2}{6}.$

(b) Prouver, pour tout entier naturel n non nul, l'inégalité : $h_n \geq (\alpha+1) \frac{n}{(n+1)^{\alpha+1}}.$

(c) Prouver l'égalité : $\lim_{\alpha \rightarrow 1^+} \sum_{n=1}^{+\infty} \frac{1}{n^\alpha} = +\infty.$

(d) Conclure que $C \geq 2$.

7. On suppose qu'il existe un réel $K > 0$ tel que, pour toute suite $(a_n)_{n \in \mathbb{N}^*}$ à termes réels strictement positifs dont la série $\sum a_n$ converge, on a $\sum_{n=1}^{+\infty} a_n \leq K \sum_{n=1}^{+\infty} h_n$.

(a) Justifier l'inégalité : $K \geq 1$.

On pourra utiliser le résultat de la question 2.

(b) Soit $N \in \mathbb{N}^*$. On pose $g_0 = 1$ et, pour tout entier naturel n non nul,

$$g_n = \begin{cases} \frac{4}{2^n} & \text{si } \exists p \in \llbracket 1, N \rrbracket \quad n = p^2 \\ \frac{1}{2^n} & \text{sinon.} \end{cases} \quad \text{et} \quad a_n = \frac{(g_n)^n}{(g_{n-1})^{n-1}}.$$

i. Calculer a_n pour $n > N^2 + 1$ et en déduire la convergence de la série $\sum a_n$.

ii. Prouver les inégalités : $\sum_{n=1}^{+\infty} g_n \leq 4$ et $\sum_{n=1}^{+\infty} a_n \geq 2N$.

(c) Établir l'égalité : $(g_n)^n = \prod_{k=1}^n a_k$.

(d) Prouver, pour tout entier naturel n non nul, l'inégalité $h_n \leq g_n$ et en déduire qu'un tel réel K n'existe pas.

Exercice 2. Premier numéro manquant

Soit $N \in \mathbb{N}^*$, $n \in \mathbb{N}^*$ avec $1 \leq n < N$ et $(X_1, X_2, \dots, X_n)$ une famille de variables aléatoires **indépendantes** toutes définies sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et toutes de loi uniforme sur l'intervalle entier $\llbracket 1, N \rrbracket$. On note, pour tout entier naturel n non nul, T l'application qui, à chaque $\omega \in \Omega$, associe

$$T(\omega) = \min \{k \in \llbracket 1, N \rrbracket; \forall i \in \llbracket 1, n \rrbracket \quad X_i(\omega) \neq k\}.$$

*Si on se représente l'expérience aléatoire consistant à tirer n fois successivement, avec remise, une boule dans une urne contenant N boules numérotées de 1 à N , la valeur $T(\omega)$ serait le numéro aléatoire du plus petit numéro qui **n'est pas** apparu dans la liste ω des numéros tirés au cours de la succession de ces n tirages.*

Partie A : Deux cas particuliers

Soit N un entier au moins égal à 4 et X, Y et Z trois variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$, **indépendantes** et toutes trois de loi uniforme sur l'intervalle entier $\llbracket 1, N \rrbracket$. On note U et V les applications qui, à chaque $\omega \in \Omega$, associent

$$U(\omega) = \min \{k \in \llbracket 1, N \rrbracket; X(\omega) \neq k \text{ et } Y(\omega) \neq k\}.$$

et

$$V(\omega) = \min \{k \in \llbracket 1, N \rrbracket; X(\omega) \neq k, Y(\omega) \neq k \text{ et } Z(\omega) \neq k\}.$$

1. (a) Quelles sont les valeurs que peut prendre la variable aléatoire U ?

- (b) Déterminer la loi de la variable aléatoire U .
 - (c) Calculer l'espérance de U .
2. (a) Calculer $\mathbf{P}(V = 1)$ et $\mathbf{P}(V = 4)$.
- (b) Établir l'égalité : $\mathbf{P}(V = 3) = \frac{6N - 12}{N^3}$.
- (c) En déduire que l'espérance de V vaut $\left(1 + \frac{1}{N}\right)^3$.

Partie B : Des égalités

On note Δ l'application qui, à tout polynôme réel $P(X)$, associe le polynôme

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

On observera que l'application Δ est linéaire.

On note, pour tout entier naturel k non nul, Δ^k l'application $\underbrace{\Delta \circ \Delta \circ \dots \circ \Delta}_{k \text{ fois } \Delta}$

et on pose $\Delta^0 = \text{id}_{\mathbb{R}[X]}$.

1. Soit $(j, n) \in \mathbb{N}^2$ avec $j \leq n$. Établir l'égalité : $\sum_{k=j}^n \binom{k}{j} = \binom{n+1}{j+1}$.
2. Soit $n \in \mathbb{N}$ et $P := \sum_{k=0}^n a_k X^k$.

Prouver, pour tout entier naturel m , l'égalité :

$$\Delta^m(P)(X) = \sum_{j=0}^m (-1)^{m-j} \binom{m}{j} P(X + j).$$

On pourra procéder par récurrence sur l'entier m .

3. (a) Soit $q \in \mathbb{N}$ et $Q(X) \in \mathbb{R}[X]$ de degré q .
Quel est, pour tout entier $r \in \llbracket 0, q \rrbracket$, le degré de $\Delta^r(Q)(X)$?
- (b) Soit $n \in \mathbb{N}$ et $Q = X^n$. Que vaut $\Delta^n(Q)(X)$?
4. Soit $n \in \mathbb{N}$ et $P := \sum_{k=0}^n a_k X^k$ un polynôme réel de degré n .

(a) Déduire des questions précédentes l'égalité :

$$a_n n! = \sum_{j=0}^n (-1)^{n-j} \binom{n}{j} P(X + j).$$

(b) En déduire l'égalité :

$$\sum_{j=0}^n (-1)^j \binom{n}{j} \left(1 - \frac{j}{N}\right)^n = \frac{n!}{N^n}.$$

On utilisera le résultat précédent pour un polynôme P bien choisi.

5. Soit $(n, N) \in \mathbb{N} \times \mathbb{N}^*$.

(a) Établir, pour tout entier $m \geq n+1$, l'égalité :

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

(b) En déduire la valeur de la somme $\sum_{j=1}^{n+1} (-1)^{j-1} \binom{n+1}{j} \left(1 - \frac{j}{N}\right)^n$.

Partie C : Le cas général

On reprend maintenant les notations du préambule.

Soit $(N, n) \in \mathbb{N}^2$ avec $1 \leq n < N$.

1. (a) Déterminer la valeur de $\mathbf{P}(T = 1)$.
- (b) Déterminer la valeur de $\mathbf{P}(T = n+1)$.
- (c) Établir l'égalité

$$\mathbf{P}(T = 2) = \sum_{i=1}^n \binom{n}{i} \left(\frac{1}{N}\right)^i \left(1 - \frac{2}{N}\right)^{n-i} \quad \text{puis l'égalité} \quad \mathbf{P}(T = 2) = \left(1 - \frac{1}{N}\right)^n - \left(1 - \frac{2}{N}\right)^n.$$

(d) Établir l'égalité : $\mathbf{P}(T = 3) = \left(1 - \frac{1}{N}\right)^n - 2\left(1 - \frac{2}{N}\right)^n + \left(1 - \frac{3}{N}\right)^n$.

2. On admet que si $A_1, A_2, \dots, A_n$ sont n événements d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ alors on dispose de la formule du crible suivante :

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{j=1}^n (-1)^{j-1} \sum_{1 \leq i_1 < i_2 < \dots < i_j \leq n} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_j}).$$

On pose, pour tout entier $p \in \llbracket 1, n \rrbracket$, $A_p = \bigcap_{i=1}^n [X_i \neq p]$.

- (a) Soit $j \in \llbracket 1, n \rrbracket$. Que vaut la probabilité de l'intersection de j événements distincts pris parmi $A_1, A_2, \dots, A_n$?
 - (b) On note E^c le complémentaire d'un événement E . Justifier, pour tout entier $k \in \llbracket 1, n \rrbracket$, l'égalité : $[T > k] = \left(\bigcup_{i=1}^k A_i\right)^c$.
 - (c) Obtenir, pour tout entier $k \in \llbracket 1, n \rrbracket$, une expression de $\mathbf{P}(T > k)$ faisant intervenir un symbole de sommation.
3. (a) Soit $k \in \llbracket 1, n+1 \rrbracket$. Établir l'égalité : $\mathbf{P}(T = k) = \sum_{i=1}^k (-1)^{i-1} \binom{k-1}{i-1} \left(1 - \frac{i}{N}\right)^n$.
 - (b) Retrouver, à l'aide du résultat précédent, les valeurs trouvées en fin de **partie A** pour les sommes

$$\sum_{j=0}^n (-1)^j \binom{n}{j} \left(1 - \frac{j+1}{N}\right)^n \quad \text{et} \quad \sum_{i=1}^{n+1} (-1)^{i-1} \binom{n+1}{i} \left(1 - \frac{i}{N}\right)^n.$$

4. Calcul de l'espérance de T

Soit $(N, n) \in \mathbb{N}^2$ avec $1 \leq n < N$.

(a) Établir, pour tout entier $k \in \llbracket 1, n \rrbracket$, l'égalité :

$$\mathbf{P}(T > k) = \left(\frac{N+1}{N}\right)^n \sum_{j=0}^k (-1)^j \binom{k}{j} \left(\frac{N+1-j}{N+1}\right)^n.$$

(b) En utilisant le résultat de la question **B-5-b)** (avec $N \leftarrow N+1$) conclure à l'égalité :

$$\mathbf{E}(T) = \left(1 + \frac{1}{N}\right)^n.$$

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2023

Épreuve à option (A) : Mathématiques

Durée : 4h

Le problème porte sur une famille de polynômes orthogonaux, appelés polynômes de Hermite, qui interviennent dans de nombreuses applications, en physique, en théorie du signal et en probabilité, par exemple.

L'énoncé est divisé en trois parties, largement indépendantes, que les candidats ne sont pas tenus de traiter dans l'ordre.

L'évaluation des copies sera étroitement liée à la rigueur des raisonnements et à une utilisation dûment justifiée du cours. Une présentation soignée sera appréciée, une présentation par trop négligée sanctionnée.

On note $\mathcal{L}_2$ l'ensemble des applications u de $\mathbb{R}$ dans $\mathbb{R}$, continues sur $\mathbb{R}$ et telles que l'intégrale généralisée $\int_{-\infty}^{+\infty} (u(x))^2 e^{-x^2} dx$ converge.

On note $\mathcal{R}$ le $\mathbb{R}$ -espace vectoriel des applications polynomiales de $\mathbb{R}$ dans $\mathbb{R}$, qu'on peut confondre avec les polynômes de $\mathbb{R}[X]$ pour en définir les coefficients et le degré (égal à $-\infty$ pour la fonction nulle).

Pour tout $n \in \mathbb{N}$, on note $\mathcal{R}_n$ le $\mathbb{R}$ -espace vectoriel des applications polynomiales de $\mathbb{R}$ dans $\mathbb{R}$, de degré inférieur ou égal à n .

Partie 1 Étude et développements en séries d'une fonction de deux variables

On note f la fonction définie sur $\mathbb{R}^2$ par :

$$\forall (x, y) \in \mathbb{R}^2, \quad f(x, y) = e^{2xy - x^2}.$$

- Justifier que f est de classe C^∞ sur $\mathbb{R}^2$ et que son seul point critique est l'origine $(0, 0)$. La fonction f admet-elle un extremum local en ce point?

- On note K l'ensemble des éléments (x, y) de $\mathbb{R}^2$ tels que $\frac{x^2}{4} + y^2 \leq 1$.

a) Justifier que f admet un maximum et un minimum sur K .

b) Soit $(x, y) \in \mathbb{R}^2$ tel que $\frac{x^2}{4} + y^2 = 1$.

Justifier qu'il existe un réel t tel que, pour cette valeur de t , $\begin{cases} x = 2 \cos t \\ y = \sin t \end{cases}$ et :

$$2xy - x^2 = 2\sqrt{2} \sin\left(2t - \frac{\pi}{4}\right) - 2.$$

c) En déduire le maximum et le minimum de f sur K , puis les points de K où ils sont atteints.

- Justifier que, pour tout $x \in \mathbb{R}$, la fonction $y \mapsto f(x, y)$ est développable en série entière sur $\mathbb{R}$ et en donner le développement en série entière.
- On note w l'application de $\mathbb{R}$ dans $\mathbb{R}$, de classe C^∞ , définie par

$$\forall x \in \mathbb{R}, \quad w(x) = e^{-x^2}.$$

Pour tout $n \in \mathbb{N}$, on note H_n l'application de $\mathbb{R}$ dans $\mathbb{R}$ définie par

$$\forall x \in \mathbb{R}, \quad H_n(x) = (-1)^n e^{x^2} w^{(n)}(x)$$

où $w^{(n)}$ désigne la dérivée n -ième de w .

En particulier : $H_0(x) = 1$.

a) Calculer, pour tout $x \in \mathbb{R}$, $H_1(x)$, $H_2(x)$, $H_3(x)$.

b) Montrer, pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$:

$$H_{n+1}(x) = 2xH_n(x) - H'_n(x).$$

c) En déduire que, pour tout $n \in \mathbb{N}$, H_n est un polynôme de degré n et trouver son coefficient dominant.

5. a) Compléter l'avant-dernière ligne du code Python suivant (en substituant une expression convenable aux trois points d'interrogation) pour que la fonction "coefH" calcule les coefficients du polynôme P_n pour la valeur de n entrée en paramètre.

```
def coefH(n):
    coeff=numpy.zeros((n+1,n+1))
    coeff[0,0]=1
    coeff[1,1]=2
    for k in range(1,n):
        coeff[k+1,0]=-coeff[k,1]
        coeff[k+1,k+1]=2**(k+1)
        for i in range(1,k+1):
            coeff[k+1,i]=2*coeff[k,i-1]- ??? # ligne à modifier
    return coeff[n,:]

coefH(4) # exemple: les coefficients de H4, égal à 16 X^4 - 48 X^2 + 12
Out: array([ 12., -0., -48.,  0.,  16.] )
```

b) Qu'obtiendrait-on pour "coefH(5)" après avoir modifié le code de la fonction "coefH" en insérant avant sa dernière ligne ("return coeff[n, :]") l'instruction suivante?

```
print(coeff)
```

6. a) Justifier que, pour tout $y \in \mathbb{R}$, la fonction $x \mapsto f(x, y)$ est développable en série entière sur $\mathbb{R}$.

b) Pour tout $(x, y) \in \mathbb{R}^2$, établir l'égalité : $f(x, y) = \sum_{n=0}^{+\infty} \frac{H_n(y)}{n!} x^n$.

Partie 2 Un produit scalaire sur $\mathcal{L}_2$

1. a) Justifier que, pour tout $(u, v) \in \mathcal{L}_2^2$, l'intégrale $\int_{-\infty}^{+\infty} u(x)v(x)e^{-x^2} dx$ converge.

On note $(.|.)$ l'application de $\mathcal{L}_2^2$ dans $\mathbb{R}$ définie par :

$$\forall (u, v) \in \mathcal{L}_2^2, \quad (u|v) = \int_{-\infty}^{+\infty} u(x)v(x)e^{-x^2} dx.$$

b) Justifier que $\mathcal{L}_2$ est un $\mathbb{R}$ -espace vectoriel et que l'application $(.|.)$ est un produit scalaire sur $\mathcal{L}_2$.

On note $\|.\|$ la norme sur $\mathcal{L}_2$ associée au produit scalaire $(.|.)$, définie par :

$$\forall u \in \mathcal{L}_2, \quad \|u\| = \sqrt{(u|u)}.$$

2. On note $\mathcal{P}$ l'ensemble des fonctions paires de $\mathcal{L}_2$ et $\mathcal{I}$ l'ensemble des fonctions impaires de $\mathcal{L}_2$.

a) Justifier l'égalité :

$$\mathcal{L}_2 = \mathcal{P} \oplus \mathcal{I} \quad . \quad (1)$$

b) Pour tout entier $n \in \mathbb{N}^*$, on note $\mathcal{P}_n = \mathcal{P} \cap \mathcal{R}_n$ et $\mathcal{I}_n = \mathcal{I} \cap \mathcal{R}_n$.

Trouver, pour tout $n \in \mathbb{N}^*$, la dimension de $\mathcal{P}_n$ et la dimension de $\mathcal{I}_n$ (on distinguera le cas où n est pair du cas où n est impair).

c) Pour tout $u \in \mathcal{I}$ et tout $n \in \mathbb{N}^*$, calculer $\min \left\{ \int_{-\infty}^{+\infty} (u(x) - P(x))^2 e^{-x^2} dx; P \in \mathcal{P}_n \right\}$.

3. On note $\mathcal{K}$ l'ensemble des fonctions de $\mathcal{L}_2$ qui sont nulles hors d'un segment inclus dans l'intervalle ouvert $]0, +\infty[$:

$$\mathcal{K} = \{v \in \mathcal{L}_2; \exists \alpha > 0, \exists A > \alpha, \forall x \notin [\alpha, A], v(x) = 0\}.$$

Soit u une fonction bornée appartenant à $\mathcal{L}_2$ et ε un nombre réel strictement positif.

a) Pour tout entier n strictement supérieur à 1, on pose :

$$v_n(x) = \begin{cases} 0 & \text{si } x < \frac{1}{n} \\ (nx - 1) u\left(\frac{2}{n}\right) & \text{si } \frac{1}{n} \leq x \leq \frac{2}{n} \\ u(x) & \text{si } \frac{2}{n} < x < n - \frac{1}{n} \\ n(n - x) u\left(n - \frac{1}{n}\right) & \text{si } n - \frac{1}{n} \leq x \leq n \\ 0 & \text{si } x > n \end{cases}.$$

Démontrer que l'intégrale $\int_0^{+\infty} (u(x) - v_n(x))^2 e^{-x^2} dx$ tend vers 0 quand n tend vers l'infini.

On pourra utiliser le théorème de convergence dominée.

b) Justifier l'existence d'un élément v de $\mathcal{K}$ tel que :

$$\int_0^{+\infty} (u(x) - v(x))^2 e^{-x^2} dx \leq \varepsilon.$$

c) Justifier que si la fonction u est paire, alors il existe un élément v de $\mathcal{K}$ tel que :

$$\|u - v_p\|^2 \leq 2\varepsilon \quad \text{où} \quad v_p(x) = \begin{cases} v(x) & \text{si } x \geq 0 \\ v(-x) & \text{si } x < 0 \end{cases}.$$

On démontre de même, et on pourra l'admettre, que si la fonction u est impaire, alors il existe un élément v de $\mathcal{K}$ tel que :

$$\|u - v_i\|^2 \leq 2\varepsilon \quad \text{où} \quad v_i(x) = \begin{cases} v(x) & \text{si } x \geq 0 \\ -v(-x) & \text{si } x < 0 \end{cases}.$$

Partie 3 Un endomorphisme autoadjoint de $\mathcal{R}$

On note g et h les endomorphismes de $\mathcal{R}$ définis par :

$$\forall P \in \mathcal{R}, \quad \begin{cases} g(P) = -P'' + 2XP' + P \\ h(P) = 2XP - P' \end{cases}.$$

Ainsi, par exemple, pour tout $P \in \mathcal{R}$ et tout $x \in \mathbb{R}$: $(h(P))(x) = 2xP(x) - P'(x)$.

1. a) Établir l'égalité : $g \circ h - h \circ g = 2h$.
b) Montrer que, pour tout $\lambda \in \mathbb{R}$ et tout $P \in \mathcal{R}$, si $g(P) = \lambda P$, alors

$$g(h(P)) = (\lambda + 2)h(P) \quad (*)$$

2. Établir, pour tout $(P, Q) \in \mathcal{R}^2$:

$$(P'|Q') = (g(P)|Q) - (P|Q).$$

3. Soit $n \in \mathbb{N}$.

a) Montrer : $\forall P \in \mathcal{R}_n, \quad g(P) \in \mathcal{R}_n$.

b) On note g_n l'endomorphisme de $\mathcal{R}_n$ défini par :

$$\forall P \in \mathcal{R}_n, \quad g_n(P) = g(P).$$

Montrer que g_n est un endomorphisme autoadjoint de $\mathcal{R}_n$, pour le produit scalaire $(\cdot|\cdot)$.

4. Les polynômes H_n utilisés dans la suite sont ceux de la partie 1.

a) Calculer, pour tout $k \in \mathbb{N}$, $h(H_k)$, et en déduire que :

$$\forall k \in \mathbb{N}, \quad g(H_k) = (2k+1)H_k.$$

b) Justifier que, pour tout $n \in \mathbb{N}^*$, la famille $(H_0, \dots, H_n)$ est une base orthogonale de $\mathcal{R}_n$.

c) En déduire, pour tout $u \in \mathcal{L}_2$, la convergence de la série $\sum_{n \geq 0} \frac{(u|H_n)^2}{\|H_n\|^2}$ et l'inégalité :

$$\sum_{n=0}^{+\infty} \frac{(u|H_n)^2}{\|H_n\|^2} \leq \int_{-\infty}^{+\infty} (u(x))^2 e^{-x^2} dx. \quad (2)$$

5. Pour tout entier $p \in \mathbb{N}^*$, on note e_p la fonction $x \mapsto e^{-px}$.

a) Établir, pour tout $n \in \mathbb{N}$ et tout réel $x > 0$, les inégalités :

$$|e^{-2x} - \sum_{k=0}^n \frac{(-1)^k}{k!} x^k e^{-x}| \leq \frac{x^{n+1}}{(n+1)!} e^{-x} \leq \frac{(n+1)^n}{n!} e^{-n-1}.$$

b) En déduire une suite de polynômes $(Q_{2,n})_{n \in \mathbb{N}}$ telle que la suite des fonctions $x \mapsto Q_{2,n}(x)e^{-x}$ converge uniformément sur $\mathbb{R}_+$ vers la fonction e_2 .

c) Démontrer par récurrence que, pour tout entier p supérieur ou égal à 2, il existe une suite de polynômes $(Q_{p,n})_{n \in \mathbb{N}}$ telle que la suite des fonctions $x \mapsto Q_{p,n}(x)e^{-x}$ converge uniformément sur $\mathbb{R}_+$ vers la fonction e_p .

On pourra utiliser les fonctions polynomiales $x \mapsto Q_{p,n}(\frac{x}{2}) \left(\sum_{k=0}^n \frac{(-1)^k (p+1)^k}{k! 2^k} x^k \right)$.

d) En déduire que, pour toute fonction polynomiale P , il existe une suite $(Q_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{R}$ telle que la suite des fonctions $x \mapsto Q_n(x)e^{-x}$ converge uniformément sur $\mathbb{R}_+$ vers la fonction $x \mapsto P(e^{-x})$.

6. Soit v une fonction appartenant à l'ensemble $\mathcal{K}$ défini dans la partie 2.

a) En utilisant le théorème de Weierstrass et le changement de variable $x = -\ln(t)$, justifier, pour tout $\varepsilon > 0$, l'existence d'un polynôme P tel que :

$$\forall x \geq 0, |\nu(x) - P(e^{-x})| \leq \varepsilon.$$

b) En déduire qu'il existe une suite $(Q_n)_{n \in \mathbb{N}}$ de polynômes (éléments de $\mathcal{R}$) telle que la suite des fonctions $x \mapsto Q_n(x)e^{-x}$ converge uniformément sur $\mathbb{R}_+$ vers la fonction v .

c) On note w la fonction définie sur $\mathbb{R}$ par

$$w(x) = \begin{cases} v(\frac{x^2}{4})e^{x^2/4} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

et, pour tout $x \in \mathbb{R}$, on pose :

$$\begin{cases} w_p(x) = w(x) + w(-x) \\ w_i(x) = w(x) - w(-x) \end{cases}.$$

Pour tout réel $\varepsilon > 0$, justifier l'existence de deux polynômes Q et R vérifiant les inégalités

$$\begin{cases} \int_{-\infty}^{+\infty} (w_p(x) - Q(\frac{x^2}{4}))^2 e^{-x^2} dx \leq \varepsilon^2 \int_{-\infty}^{+\infty} e^{-x^2/2} dx \\ \int_{-\infty}^{+\infty} (w_i(x) - xR(\frac{x^2}{4}))^2 e^{-x^2} dx \leq \varepsilon^2 \int_{-\infty}^{+\infty} x^2 e^{-x^2/2} dx \end{cases}.$$

Pour parvenir à la seconde inégalité, on pourra, en lieu et place de v , utiliser la fonction

$$t \mapsto \begin{cases} \frac{w(2\sqrt{t})}{2\sqrt{t}} e^{-t} & \text{si } t > 0 \\ 0 & \text{sinon} \end{cases}.$$

d) Soit $\eta > 0$.

Déduire des résultats précédents que :

- pour toute fonction paire w_p telle que la fonction

$$x \mapsto \begin{cases} w_p(x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

appartient à $\mathcal{K}$, il existe un polynôme pair P_p tel que

$$\|w_p - P_p\| \leq \eta$$

- pour toute fonction impaire w_i telle que la fonction

$$x \mapsto \begin{cases} w_i(x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

appartient à $\mathcal{K}$, il existe un polynôme impair P_i tel que

$$\|w_i - P_i\| \leq \eta.$$

7. Soit u une fonction continue et bornée sur $\mathbb{R}$.

a) En s'appuyant sur les derniers résultats de la partie 2 et sur ce qui précède, démontrer qu'il existe une suite d'éléments de $\mathcal{R}$ qui converge vers u pour la norme $\|\cdot\|$.

b) Justifier l'égalité :

$$\lim_{n \rightarrow +\infty} \left\| u - \sum_{k=0}^n \frac{(u|H_k)}{\|H_k\|^2} H_k \right\| = 0.$$

c) Justifier que l'inégalité (2) démontrée en question 4 est en fait une égalité :

$$\sum_{n=0}^{+\infty} \frac{(u|H_n)^2}{\|H_n\|^2} = \int_{-\infty}^{+\infty} (u(x))^2 e^{-x^2} dx. \quad (3)$$

d) Ces propriétés restent-elles vraies si u appartient à $\mathcal{L}_2$ sans être bornée?

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2023

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet comporte un problème et un exercice qui peuvent être traités indépendamment.
La notation tiendra largement compte de la clarté et de la précision des réponses.

Quelques notations et rappels

Les notations et rappels ci-dessous vous seront utiles pour la résolution du problème et de l'exercice. Pensez à vous y référer.

- Toutes les variables aléatoires considérées dans ce sujet seront définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
- Soient $(A_i)_{i \in \mathbb{N}}$ une suite d'éléments de $\mathcal{A}$. On a

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \mathbb{P}(A_i).$$

- Soit E un ensemble quelconque. Pour tout $A \subset E$ et $B \subset E$, on note

$$A \setminus B := \{x \in E \mid x \in A \text{ et } x \notin B\}.$$

- **[Formule des espérances totales]** Soient Y et N deux variables aléatoires. Si la variable aléatoire N est à valeurs dans $\mathbb{N}$ et si Y est intégrable (i.e., $\mathbb{E}(|Y|) < \infty$) alors

$$\mathbb{E}(Y) = \sum_{k \in \mathbb{N}} \mathbb{P}([N = k]) \mathbb{E}(Y \mid N = k).$$

- La somme de n variables aléatoires indépendantes de loi de Poisson de paramètres $\lambda_1 > 0, \dots, \lambda_n > 0$ suit également une loi de Poisson de paramètre $\lambda_1 + \dots + \lambda_n$.
- **[Théorème des valeurs intermédiaires]** Soit $g : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Pour tout x tel que $\min(g(a), g(b)) \leq x \leq \max(g(a), g(b))$, il existe au moins un point $c \in [a, b]$ tel que $g(c) = x$.
- Une fonction φ définie sur un intervalle $I \subset \mathbb{R}$ et à valeurs dans $\mathbb{R}$ est strictement convexe si pour tout $x \neq y \in I$ et $\lambda \in]0, 1[$, on a $\varphi(\lambda x + (1 - \lambda)y) < \lambda \varphi(x) + (1 - \lambda)\varphi(y)$.
- Le développement en série entière de l'exponentielle est donné pour tout $x \in \mathbb{R}$ par

$$e^x = \sum_{i \in \mathbb{N}} \frac{x^i}{i!}.$$

Problème – Processus de Galton-Watson

On souhaite dans ce problème modéliser l'évolution de la taille d'une population de bactéries. On adopte pour se faire le modèle suivant.

- La population initiale (génération 0) est composée d'une seule bactérie.
- Cette bactérie va engendrer un nombre aléatoire de nouvelles bactéries. Ce nombre est modélisé par une variable aléatoire $X_{1,0}$ à valeurs dans l'ensemble des entiers $\mathbb{N}$. La taille de la population de la génération 1 est donc $Z_1 := X_{1,0}$.
- Les Z_1 bactéries de la première génération vont à leur tour engendrer de nouvelles bactéries. Pour $i = 1, \dots, Z_1$, on notera $X_{i,1}$ la variable aléatoire (à valeurs dans $\mathbb{N}$) modélisant le nombre de descendants de la i -ème bactérie. Ainsi, la taille de la population de la génération 2 est $Z_2 = 0$ si $Z_1 = 0$ et

$$Z_2 := \sum_{i=1}^{Z_1} X_{i,1} \text{ si } Z_1 > 0.$$

- De manière plus générale, pour tout $n \in \mathbb{N}$, on notera Z_n la variable aléatoire correspondant au nombre d'individus de la génération n . On a donc $Z_0 = 1$ et

$$Z_n := \begin{cases} 0 & \text{si } Z_{n-1} = 0, \\ \sum_{i=1}^{Z_{n-1}} X_{i,n-1} & \text{si } Z_{n-1} > 0, \end{cases} \quad (1)$$

où pour tout $i = 1, \dots, Z_{n-1}$, la variable aléatoire $X_{i,n-1}$ est le nombre de descendants de la i -ème bactérie de la génération $n-1$.

- On supposera pour finir que les variables aléatoires $\{X_{i,n}; i \in \mathbb{N} \setminus \{0\}, n \in \mathbb{N}\}$ sont indépendantes et de même loi qu'une variable aléatoire intégrable X prenant ses valeurs dans $\mathbb{N}$. La loi de X est déterminée par les valeurs $\{p_k := \mathbb{P}([X = k]); k \in \mathbb{N}\}$ et on a donc

$$\mu := \mathbb{E}(X) = \sum_{k \in \mathbb{N}} k p_k < \infty.$$

Le processus $\{Z_n; n \in \mathbb{N}\}$ ainsi construit est connu sous le nom de *processus de Galton-Watson*. L'objectif de ce problème est d'étudier la probabilité d'extinction

$$\bar{\pi} := \mathbb{P}([\exists n \in \mathbb{N} \text{ tel que } Z_n = 0]),$$

de la population de bactéries.

Nous allons dans un premier temps calculer la probabilité d'extinction dans des cas simples.

- 1) Quelle est la valeur de $\bar{\pi}$ lorsque $p_0 = \mathbb{P}([X = 0]) = 0$? Même question lorsque $p_0 = 1$.

Dans toute la suite de ce problème, pour tout $n \in \mathbb{N}$, on notera π_n la probabilité que la population soit éteinte à la génération n . On a donc $\pi_n = \mathbb{P}([Z_n = 0])$.

- 2) Montrer que $(\pi_n)_{n \in \mathbb{N}}$ est une suite croissante et en déduire qu'elle converge vers une limite finie.
- 3) On introduit les événements $B_1 = [Z_1 = 0]$ et $\{B_i := [Z_i = 0] \setminus [Z_{i-1} = 0]; i \geq 2\}$. Montrer que pour $n \in \mathbb{N} \setminus \{0\}$,

$$\pi_n = \sum_{i=1}^n \mathbb{P}(B_i).$$

- 4) En déduire que $\bar{\pi} = \lim \pi_n$ lorsque $n \rightarrow \infty$.
- 5) Pour cette question, on suppose que $p_0 + p_1 = 1$ avec $p_0 \in]0, 1[$.
 - 5.1) Montrer que pour tout $n \in \mathbb{N}$, la variable aléatoire Z_n prend ses valeurs dans l'ensemble $\{0, 1\}$.
 - 5.2) Donner, en fonction de n et p_1 , la valeur de π_n .
 - 5.3) En déduire la valeur de $\bar{\pi}$.

Pour déterminer la probabilité d'extinction, on doit étudier la *fonction génératrice des probabilités* de la variable aléatoire X modélisant le nombre de descendants d'une bactérie. Cette fonction est définie pour tout $s \in [0, 1]$ par $G_X(s) := \mathbb{E}(s^X)$.

- 6) Donner, en fonction des probabilités $\{p_k; k \in \mathbb{N}\}$, les expressions de la suite positive $(a_k)_{k \in \mathbb{N}}$ et des fonctions positives f_k telles que pour tout $s \in [0, 1]$,

$$G_X(s) = \sum_{k \in \mathbb{N}} a_k f_k(s).$$

- 7) Montrer que la série ci-dessus est absolument convergente.
- 8) Dans le cas particulier où X suit une loi de Poisson de paramètre $\lambda > 0$, donner l'expression de $G_X(s)$ (vous devrez expliciter les calculs faits pour trouver cette expression).
- 9) Donner l'expression de $G_X(0)$ en fonction des $\{p_k; k \in \mathbb{N}\}$ (on utilisera la convention $0^0 = 1$). Quelle est la valeur de $G_X(1)$?

On admettra que la fonction G_X est indéfiniment dérivable sur $]0, 1[$. Pour tout $j \in \mathbb{N} \setminus \{0\}$, la dérivée d'ordre j est donnée par

$$G_X^{(j)}(s) = \sum_{k \in \mathbb{N}} a_k f_k^{(j)}(s), s \in]0, 1[.$$

On admettra enfin que les fonctions G_X et $G_X^{(1)}$ sont continues à droite en 0 et à gauche en 1.

- 10) Que représente la valeur $G_X^{(1)}(1)$?

- 11) Que faut-il imposer sur la valeur de p_0 pour assurer que G_X est strictement croissante sur $]0, 1[$?
- 12) Montrer que si $p_0 + p_1 < 1$ alors G_X est strictement convexe.

On souhaite à présent trouver l'expression de la fonction génératrice des probabilités G_{Z_n} de la variable aléatoire Z_n définie par l'équation 1. Remarquons tout d'abord que puisque $Z_1 = X_{1,0}$, la variable aléatoire Z_1 a même loi que X et donc $G_{Z_1}(s) = G_X(s)$ pour tout $s \in [0, 1]$.

- 13) Montrer que pour tout $n \geq 2$, la variable aléatoire Z_{n-1} est indépendante des variables aléatoires $\{X_{i,n-1}; i \geq 1\}$.
- 14) En utilisant la question précédente ainsi que la formule des espérances totales, montrer que pour tout $n \geq 1$ et $s \in [0, 1]$ on a $G_{Z_n}(s) = G_{Z_{n-1}}(G_X(s))$.
- 15) Montrer par récurrence que pour $n \geq 1$ et $s \in [0, 1]$ on a $G_{Z_n}(s) = G_X(G_{Z_{n-1}}(s))$. En déduire que $\pi_n = G_X(\pi_{n-1})$.
- 16) En utilisant la question 4), montrer que $\bar{\pi} = G_X(\bar{\pi})$.

On s'intéresse à présent à l'équation

$$G_X(s) = s, \quad s \in [0, 1]. \quad (2)$$

- 17) Donner une solution évidente de l'équation (2).
- 18) Montrer que si $u \in [0, 1]$ est solution de (2) alors $\pi_n \leq u$ pour tout $n \in \mathbb{N}$. (Aide : vous pourrez démontrer ce résultat par récurrence.)
- 19) En déduire que $\bar{\pi}$ est la plus petite solution de (2).
- 20) Pour cette question on suppose que $p_0 + p_1 < 1$ et qu'il existe au moins un point $x_0 \in]0, 1[$ tel que $G_X(x_0) < x_0$.
 - 20.1) Montrer qu'il existe au moins un point $x_1 \in [0, 1[$ solution de (2). (Aide : appliquer le théorème des valeurs intermédiaires à la fonction définie sur $[0, 1]$ par $g(s) := G_X(s) - s$).
 - 20.2) Montrer que x_1 est l'unique élément de $[0, 1[$ solution de (2). (Aide : montrer que si G_X est une fonction strictement convexe alors pour tout $x_2 \in]x_1, 1[$ on a $G_X(x_2) < x_2$).
 - 20.3) Montrer que si $G_X^{(1)}(1) \leq 1$ alors pour tout $s \in [0, 1[$ on a $G_X(s) > s$. (Aide : utiliser le fait que le graphe d'une fonction strictement convexe est au dessus de toutes ses tangentes).
 - 20.4) Démontrer que la réciproque du résultat établi dans la question précédente est vraie. Vous pourrez utiliser le fait que

$$G_X^{(1)}(1) = \lim_{\varepsilon \rightarrow 0+} \frac{G_X(1) - G_X(1 - \varepsilon)}{\varepsilon}.$$

- 21) Utiliser les résultats de la question 20) pour montrer le résultat suivant.

Proposition – i) Si $p_0 + p_1 < 1$ et $\mu := \mathbb{E}(X) \in [0, 1]$ alors $\bar{\pi} = 1$.

ii) Si $p_0 + p_1 < 1$ et $\mu > 1$ alors $\bar{\pi} \in [0, 1[$ est l'unique solution sur $[0, 1[$ de l'équation $G_X(s) = s$.

On s'intéresse à présent au nombre de générations de bactéries observées avant l'extinction de la population. Plus précisément, on introduit la variable aléatoire τ à valeurs dans $\mathbb{N} \setminus \{0\}$ et telle que $Z_\tau > 0$ et $Z_{\tau+1} = 0$.

22) Montrer que $\mathbb{P}(\tau \geq n) = \mathbb{P}(Z_n > 0)$.

23) En utilisant l'indépendance entre Z_n et les variables aléatoires $(X_{i,n}, i \in \mathbb{N})$, montrer que pour tout $k \geq 1$,

$$\mathbb{E}(Z_{n+1} \mid Z_n = k) = k\mu.$$

24) Montrer par récurrence que pour tout $n \geq 1$ on a $\mathbb{E}(Z_n) = \mu^n = [\mathbb{E}(X)]^n$ (Aide : utiliser la formule des espérances totales).

25) En déduire que $\mathbb{P}(\tau \geq n) \leq \mu^n$.

Exercice – Approximation d'une loi binomiale par une loi de Poisson

L'objectif de cet exercice est de démontrer un résultat sur l'approximation d'une loi binomiale par une loi de Poisson. Ce résultat, établi par Lucien Le Cam, s'énonce ainsi.

Pour $n \in \mathbb{N} \setminus \{0\}$, soient S_n et T_n deux variables aléatoires distribuées respectivement selon une loi de poisson de paramètres np avec $p \in]0, 1[$ et une loi binomiale de paramètres n et p . Pour tout $n \in \mathbb{N} \setminus \{0\}$ et $p \in]0, 1[$,

$$\sum_{k \in \mathbb{N}} |\mathbb{P}([S_n = k]) - \mathbb{P}([T_n = k])| \leq 4np^2. \quad (3)$$

En particulier, si $p = p_n = \lambda/n$ avec $\lambda > 0$, l'inégalité ci-dessus implique que

$$\lim_{n \rightarrow \infty} \sum_{k \in \mathbb{N}} \left| \mathbb{P}([T_n = k]) - e^{-\lambda} \frac{\lambda^k}{k!} \right| = 0.$$

En d'autres termes, la loi binomiale de paramètres n et λ/n peut être approchée pour n suffisamment grand par une loi de Poisson de paramètre λ .

Les questions ci-dessus reprennent le schéma de la démonstration de L. Le Cam. On intro-

duit tout d'abord le couple aléatoire $(X, Y) : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{N}^2, \mathcal{P}(\mathbb{N}^2))$ tel que, pour $p \in]0, 1[$,

$$\mathbb{P}([X = i] \cap [Y = j]) = \begin{cases} e^{-p} - p + pe^{-p} & \text{si } i = j = 0, \\ p - pe^{-p} & \text{si } i = 0 \text{ et } j = 1, \\ pe^{-p} & \text{si } i = j = 1, \\ e^{-p} p^i / i! & \text{si } i \geq 2 \text{ et } j = 0 \end{cases}$$

- 1) Donner l'expression, en fonction de p , de la probabilité $\mathbb{P}([X \geq 2] \cap [Y = 0])$.
- 2) Déterminer l'ensemble $E \subset \mathbb{N}^2$ des valeurs prises par le couple aléatoire (X, Y) .
- 3) Pour tout $k \in \mathbb{N}$, donner en fonction de p l'expression de la probabilité $\mathbb{P}([X = k])$.
Quelle est la loi suivie par la variable aléatoire X ?
- 4) Donner les expressions des probabilités $\mathbb{P}([Y = 0])$ et $\mathbb{P}([Y = 1])$. Quelle est la loi suivie par la variable aléatoire Y ?
- 5) Justifier le fait que $\mathbb{P}([X = Y]) = \mathbb{P}([X = 0] \cap [Y = 0]) + \mathbb{P}([X = 1] \cap [Y = 1])$.
- 6) Montrer que pour tout $x \geq 0$ on a l'inégalité $e^{-x} \geq 1 - x$.
- 7) Dédire des questions 5) et 6) que $\mathbb{P}([X = Y]) \geq 1 - 2p^2$.

On introduit à présent les $n \in \mathbb{N} \setminus \{0\}$ couples aléatoires $\{(X_i, Y_i), i = 1, \dots, n\}$ que l'on suppose indépendants et de même loi que le couple aléatoire (X, Y) introduit en début d'exercice. On pose

$$S_n := \sum_{i=1}^n X_i \text{ et } T_n := \sum_{i=1}^n Y_i.$$

- 8) Quelle est la loi suivie par la variable aléatoire S_n ? Même question pour la variable aléatoire T_n .
- 9) En utilisant le fait que si $X_i = Y_i$ pour tout $i = 1, \dots, n$ alors $S_n = T_n$, montrer que

$$[S_n \neq T_n] \subset \bigcup_{i=1}^n [X_i \neq Y_i].$$

- 10) En déduire que $\mathbb{P}([S_n \neq T_n]) \leq 2np^2$.
- 11) Montrer que $\mathbb{P}([S_n \in A]) \leq \mathbb{P}([T_n \in A]) + \mathbb{P}([S_n \neq T_n])$ pour tout sous-ensemble A de $\mathbb{N}$.
(Aide : exprimer la probabilité $\mathbb{P}([S_n \in A])$ en utilisant la formule des probabilités totales avec la partition de Ω suivante : $\{[S_n = T_n], [S_n \neq T_n]\}$).
- 12) En déduire que pour tout $A \subset \mathbb{N}$, on a $\mathbb{P}([S_n \in A]) - \mathbb{P}([T_n \in A]) \leq 2np^2$.
- 13) Montrer que l'on a également $\mathbb{P}([T_n \in A]) - \mathbb{P}([S_n \in A]) \leq 2np^2$.

On introduit l'ensemble $A_n := \{k \in \mathbb{N} \mid \mathbb{P}([S_n = k]) > \mathbb{P}([T_n = k])\} \subset \mathbb{N}$.

14) Montrer que

$$\begin{aligned} \sum_{k \in \mathbb{N}} |\mathbb{P}([S_n = k]) - \mathbb{P}([T_n = k])| &= \mathbb{P}([S_n \in A_n]) - \mathbb{P}([T_n \in A_n]) \\ &+ \mathbb{P}([T_n \in \mathbb{N} \setminus A_n]) - \mathbb{P}([S_n \in \mathbb{N} \setminus A_n]). \end{aligned}$$

15) Dédurre l'inégalité (3) des questions précédentes.

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2024

Épreuve de mathématiques

Durée : 4h

L'épreuve est constituée de trois exercices indépendants.

Rappel : Pour toutes les épreuves, les appareils électroniques (calculatrices, téléphones, traductrices), les casques, et les dictionnaires, ainsi que tout document, sont interdits.

Exercice 1. Probabilités.

Si X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans $\mathbb{N}$, on note G_X sa fonction génératrice, définie pour $t \in \mathbb{R}$ par

$$G_X(t) = \sum_{n=0}^{\infty} \mathbf{P}(X = n) t^n,$$

lorsque la série converge.

1. (a) Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbf{P})$, à valeurs dans $\mathbb{N}$ et indépendantes. Déterminer la loi de $X + Y$ en fonction de celles de X et Y et en déduire une relation entre G_{X+Y} , G_X et G_Y .
- (b) Généraliser le résultat précédent à la somme de n variables aléatoires mutuellement indépendantes.
2. On veut montrer qu'il n'est pas possible de construire un dé faussé à 6 faces (numérotées de 1 à 6) tel que si on le lance deux fois, la somme des deux résultats obtenus suive une loi uniforme sur $\{2, 3, 4, \dots, 12\}$.
 - (a) Soit U la loi uniforme sur $\{2, \dots, 12\}$. Déterminer sa fonction génératrice G_U et déterminer toutes les racines réelles de G_U ainsi que leur multiplicité.
 - (b) Supposons que l'on ait un dé à 6 faces tel que la probabilité d'apparition de la face $i \in \{1, \dots, 6\}$ soit égale à p_i , avec $p_i \in]0, 1[$. Soit X la loi d'apparition de la face i . Montrer que G_X admet une racine non-nulle.
 - (c) Conclure.
3. Un commerçant reçoit dans son magasin N clients par jour, où N est une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda > 0$. Chaque client repart avec un produit avec une probabilité $p \in [0, 1]$. On supposera que les comportements de deux clients distincts sont indépendants. Soit Y la variable aléatoire égale au nombre de produits vendus par jour.
 - (a) Préciser la fonction caractéristique de N .
 - (b) Déterminer la fonction caractéristique de Y et montrer que Y suit une loi de Poisson dont on précisera le paramètre.

◇

Exercice 2. autour de la constante d'Euler.

Dans ce problème on étudie plusieurs propriétés et expressions de la constante γ d'Euler. Les parties 2 et 3 sont indépendantes et utilisent les résultats de la partie 1.

On rappelle l'énoncé du théorème de convergence dominée :

Théorème. Soit (f_n) une suite de fonctions continues par morceaux qui converge simplement vers une fonction f continue par morceaux sur un intervalle I . On suppose qu'il existe une fonction positive g telle que l'intégrale $\int_I g(x)dx$ converge et pour tout n on a $|f_n| \leq g$. Alors les intégrales $\int_I f_n(x)dx$ et $\int_I f(x)dx$ convergent absolument et $\lim_{n \rightarrow \infty} \int_I f_n(x)dx = \int_I f(x)dx$.

Partie 1 : Préambule

Pour tout $n \geq 1$, on note

$$H_n = \sum_{k=1}^n \frac{1}{k}$$

la n^{e} somme partielle de la série harmonique.

1. Montrer que pour tout $n \geq 1$ on a l'inégalité $H_{2n} - H_n \geq \frac{1}{2}$, et utiliser cette inégalité pour montrer que la série harmonique $\sum_{n \geq 1} \frac{1}{n}$ diverge.
2. (a) Montrer qu'il existe une constante C telle que pour tout $n \geq 1$, on a

$$\left| \ln \left(1 + \frac{1}{n} \right) - \frac{1}{n} \right| \leq \frac{C}{n^2}.$$

- (b) En déduire que la série $\sum_{n \geq 1} \left(\ln \left(\frac{n+1}{n} \right) - \frac{1}{n} \right)$ converge, puis qu'il existe une constante γ telle que

$$H_n = \ln n + \gamma + o(1) \text{ quand } n \rightarrow \infty.$$

Cette constante γ s'appelle la constante d'Euler, dont nous allons étudier certaines propriétés dans les parties suivantes.

Partie 2 : Une expression intégrale de la constante d'Euler.

L'objet de cette partie est d'établir la formule

$$(1) \quad \int_0^{\infty} \ln(t) e^{-t} dt = -\gamma.$$

3. Montrer que les intégrales impropres $\int_0^1 \ln(t) e^{-t} dt$ et $\int_1^{\infty} \ln(t) e^{-t} dt$ sont convergentes.
4. (a) Montrer que pour tout $u \in]-1, \infty[$ on a $\ln(1+u) \leq u$
 (b) En déduire que pour tout entier $n \geq 2$ et tout $t \in [0, n]$ on a

$$0 \leq \left(1 - \frac{t}{n} \right)^{n-1} \leq e \cdot e^{-t}.$$

5. Montrer que

$$\lim_{n \rightarrow \infty} \int_0^n \ln(t) \left(1 - \frac{t}{n} \right)^{n-1} dt = \int_0^{\infty} \ln(t) e^{-t} dt.$$

6. Montrer successivement les égalités

$$\int_0^n \ln(t) \left(1 - \frac{t}{n} \right)^{n-1} dt = \ln n + \int_0^1 n \ln(u) (1-u)^{n-1} du = \ln n + \int_0^1 \frac{(1-u)^n - 1}{u} du$$

7. En déduire la formule (1)

Partie 3 : Lien avec la fonction ζ .

Pour tout $s > 1$ on pose $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ (qui se prononce "dzêta").

8. Montrer que pour tout $s > 1$, $\zeta(s)$ est bien définie et vérifie

$$\int_1^\infty \frac{dt}{t^s} \leq \zeta(s) \leq 1 + \int_1^\infty \frac{dt}{t^s}.$$

En déduire un équivalent de $\zeta(s)$ lorsque $s \rightarrow 1^+$.

9. (a) Montrer que pour $s > 1$ la série

$$\sum_{n \geq 1} n \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right)$$

converge.

(b) En évaluant ses sommes partielles, montrer que sa somme vaut $\zeta(s)$.

(c) En déduire la formule

$$\zeta(s) = s \int_1^\infty \frac{E(t)}{t^{s+1}} dt$$

où $E(\cdot)$ désigne la fonction partie entière.

10. Démontrer que pour tout entier $N \geq 2$, on a

$$\int_1^N \frac{E(t) - t}{t^2} dt = H_N - \ln(N) - 1,$$

et en déduire la formule

$$\int_1^\infty \frac{E(t) - t}{t^2} dt = \gamma - 1.$$

11. En justifiant l'égalité $\zeta(s) - \frac{s}{s-1} = s \int_1^\infty \frac{E(t) - t}{t^{s+1}} dt$, conclure que

$$\zeta(s) = \frac{1}{s-1} + \gamma + o(1) \text{ lorsque } s \rightarrow 1^+.$$

◇

Exercice 3. Sous-espaces de matrices nilpotentes.

Soit $n \in \mathbb{N}^*$. On note $M_n(\mathbb{R})$ le $\mathbb{R}$ -espace vectoriel des matrices $n \times n$ à coefficients réels. On rappelle qu'une matrice $N \in M_n(\mathbb{R})$ est dite nilpotente s'il existe $k \geq 1$ tel que $N^k = 0$, et on note $N_n \subset M_n(\mathbb{R})$ le sous-ensemble des matrices nilpotentes. On identifiera $\mathbb{R}^n$ à l'espace $M_{n,1}(\mathbb{R})$ des vecteurs colonnes de taille n , et une matrice $M \in M_n(\mathbb{R})$ à l'endomorphisme de $\mathbb{R}^n$ qui lui est canoniquement associé. On note M^T la transposée d'une matrice M .

L'objet de ce problème est de démontrer le théorème suivant, dû à Gerstenhaber.

Théorème. Si V est un sous-espace vectoriel de $M_n(\mathbb{R})$ constitué de matrices nilpotentes, alors

$$\dim(V) \leq \frac{n(n-1)}{2}.$$

On note T_n^+ (resp. T_n^-) l'ensemble des matrices triangulaires supérieures (resp. inférieures) et T_n^{++} l'ensemble des matrices triangulaires supérieures strictes, c'est-à-dire n'ayant que des 0 sur la diagonale.

1. Montrer que T_n^+ , T_n^- et T_n^{++} sont des sous-espaces vectoriels (on pourra se contenter de rédiger les détails pour T_n^+) et préciser leur dimension.
2. Montrer que $T_n^- \oplus T_n^{++} = M_n(\mathbb{R})$.
3. Pour $1 \leq k \leq n$, on note $E_k = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n; \forall i \geq k, x_i = 0\}$. On note également $E_{n+1} = \mathbb{R}^n$. Soit $A \in T_n^{++}$ et u l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à A .
 - (a) Montrer que pour tout $1 \leq k \leq n+1$, on a $u(E_k) \subset E_{k-1}$.
 - (b) En déduire que A est nilpotente. On a ainsi montré que $T_n^{++} \subset N_n$.
 - (c) Montrer que $T_n^+ \cap N_n = T_n^{++}$.
4. N_n est-il un sous-espace vectoriel de $M_n(\mathbb{R})$?
5. Montrer que si $A \in N_n$, alors $\text{Tr}(A) = 0$. (*Indication : on pourra étudier les valeurs propres complexes de A .*)
6. Soient A et B dans $M_n(\mathbb{R})$. Montrer que si A , B et $A+B$ sont nilpotentes alors $\text{Tr}(AB) = 0$.

Dans la suite du problème on fixe un sous-espace vectoriel V de $M_n(\mathbb{R})$ contenu dans N_n .

On note π la projection de $M_n(\mathbb{R})$ sur T_n^{++} parallèlement à T_n^- (qui est bien définie grâce à la question 2). On note également τ l'endomorphisme de $M_n(\mathbb{R})$ défini par la transposition, i.e. pour tout $M \in M_n(\mathbb{R})$, $\tau(M) = M^T$.

7.
 - (a) Montrer que $\dim(V) = \dim(V \cap T_n^-) + \dim(\pi(V))$.
 - (b) En déduire que $\dim(V) = \dim(\tau(V \cap T_n^-)) + \dim(\pi(V))$.
8. Montrer que $\tau(V \cap T_n^-)$ est inclus dans T_n^{++} .
9. Soit $A \in \tau(V \cap T_n^-) \cap \pi(V)$ et on fixe $N \in V$ telle que $A = \pi(N)$.
 - (a) Montrer que $A^T + N$ appartient à V . En déduire que $\text{Tr}(A^T N) = 0$.
 - (b) Montrer que $N - A$ appartient à T_n^- et en déduire que $A(N^T - A^T) \in T_n^{++}$.
 - (c) Montrer que $\text{Tr}(AA^T) = 0$.
 - (d) Conclure que $A = 0$.
10. En utilisant les résultats des questions précédentes, démontrer le théorème de Gershtenhaber.

◇

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2024

Épreuve à option (A) : Mathématiques

Durée : 4h

Rappel : Pour toutes les épreuves, les appareils électroniques (calculatrices, téléphones, traductrices), les casques, et les dictionnaires, ainsi que tout document, sont interdits.

Ce problème est consacré à des questions d'approximation uniforme des fonctions sur un intervalle.

L'espace des fonctions continues sur un intervalle $[a, b]$ est noté $\mathcal{C}([a, b])$.

La transposée d'une matrice $M \in M_n(\mathbb{C})$ est notée $M^\top$ et son rang est noté $\text{rg}(M)$. L'espace vectoriel engendré par une partie A d'un espace vectoriel V est noté $\text{Vect}(A)$. Si $(E, \|\cdot\|)$ est un espace vectoriel normé on note $d(x, A) = \inf\{\|x - a\|, a \in A\}$ la distance d'un point $x \in E$ à une partie A de E .

Le problème est composé de quatre parties. Les trois premières parties sont indépendantes, et la quatrième utilise les résultats établis dans les parties précédentes. On attachera la plus grande importance au soin, à la clarté et à la précision de la rédaction.

Partie 1 : Déterminant de Cauchy.

On fixe un entier $n \geq 1$ et $a_1, \dots, a_n, b_1, \dots, b_n$ des nombres complexes tels que $a_i + b_i \neq 0$ pour tout $1 \leq i \leq n$. L'objet de cette partie est de calculer le déterminant de Cauchy

$$D_n = \det \left(\frac{1}{a_i + b_j} \right)_{1 \leq i, j \leq n}.$$

Pour cela on suppose dans un premier temps que les (a_i) sont distincts, de même que les (b_i) , et on introduit la fraction rationnelle

$$F(X) = \begin{vmatrix} \frac{1}{a_1 + b_1} & \frac{1}{a_1 + b_2} & \cdots & \frac{1}{a_1 + b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{n-1} + b_1} & \frac{1}{a_{n-1} + b_2} & \cdots & \frac{1}{a_{n-1} + b_n} \\ \frac{1}{X + b_1} & \frac{1}{X + b_2} & \cdots & \frac{1}{X + b_n} \end{vmatrix}$$

1. Montrer qu'il existe un nombre complexe λ tel que $F(X)$ s'écrive sous la forme

$$F(X) = \lambda \frac{(X - a_1)(X - a_2) \cdots (X - a_{n-1})}{(X + b_1)(X + b_2) \cdots (X + b_n)}.$$

2. On pose $G(X) = (X + b_n)F(X)$. Montrer que $G(-b_n) = D_{n-1}$ et en déduire une relation entre D_n et D_{n-1} .

3. Conclure que

$$D_n = \frac{\prod_{1 \leq i < j \leq n} (a_j - a_i) \prod_{1 \leq i < j \leq n} (b_j - b_i)}{\prod_{1 \leq i, j \leq n} (a_i + b_j)},$$

dans le cas où les (a_i) et les (b_i) sont distincts, puis dans le cas général.

Partie 2 : Matrices de Gram.

Dans cette partie on fixe un espace vectoriel E sur $\mathbb{R}$ muni d'un produit scalaire noté $\langle \cdot | \cdot \rangle$ (qui munit donc E d'une structure d'espace préhilbertien). La norme associée est notée $\|\cdot\|$. Si $(x_1, \dots, x_n)$ est un n -uplet de vecteurs de E on définit la *matrice de Gram* de $(x_1, \dots, x_n)$ par

$$\text{Gram}(x_1, \dots, x_n) = (\langle x_i | x_j \rangle)_{1 \leq i, j \leq n}.$$

On pose également $V = \text{Vect}(x_1, \dots, x_n)$ et $m = \dim(V)$.

4. Soit $(e_1, \dots, e_m)$ une base orthonormale de V et M la matrice de taille $m \times n$ définie par $M = (\langle e_i | x_j \rangle)_{1 \leq i \leq m, 1 \leq j \leq n}$.

(a) Montrer que

$$\text{Gram}(x_1, \dots, x_n) = M^T M.$$

(b) Montrer que $\text{rg}(\text{Gram}(x_1, \dots, x_n)) = \dim V$.

(c) Montrer que toute valeur propre réelle de $M^T M$ est positive ou nulle.

(d) En déduire que $\det(\text{Gram}(x_1, \dots, x_n)) \geq 0$ avec égalité si et seulement si $(x_1, \dots, x_n)$ est une famille liée.

5. On suppose que la famille $(x_1, \dots, x_n)$ est libre. Soit $x \in E$ et $p(x)$ le projeté orthogonal de x sur V . Montrer que

$$d(x, V) = \sqrt{\frac{\det(\text{Gram}(x, x_1, \dots, x_n))}{\det(\text{Gram}(x_1, \dots, x_n))}}.$$

(on pourra considérer le déterminant $\det(\text{Gram}(x - p(x), x_1, \dots, x_n))$)

Partie 3 : Polynômes de Bernstein et théorème de Weierstrass

Dans cette partie on s'intéresse au théorème d'approximation de Weierstrass (1885) : si $[a, b]$ est un intervalle compact de $\mathbb{R}$, toute fonction continue de $[a, b]$ dans $\mathbb{R}$ est limite uniforme sur $[a, b]$ d'une suite de polynômes.

Pour tous $n \geq 0$ et $0 \leq k \leq n$ on pose $B_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$. Cette famille s'appelle la famille des *polynômes de Bernstein*. On note $\mathcal{P}_n$ l'espace vectoriel des polynômes réels de degré au plus n .

Pour $x \in [0, 1]$ et $n \in \mathbb{N}$, on considère une suite $(X_i)_{i=1, \dots, n}$ de variables aléatoires de Bernoulli indépendantes de paramètre x : $\mathbb{P}(X_i = 1) = x$, $\mathbb{P}(X_i = 0) = 1 - x$. On fixe $f \in \mathcal{C}([0, 1])$ et on pose

$$B_n(f)(x) = \mathbb{E} \left[f\left(\frac{S_n}{n}\right) \right] \text{ où } S_n = \sum_{i=1}^n X_i.$$

Pour $f \in \mathcal{C}([0, 1])$, on définit son module de continuité ω par

$$\omega(\delta) = \sup \{ |f(x) - f(y)|, (x, y) \in [0, 1]^2, |x - y| \leq \delta \}.$$

6. Montrer que le module de continuité ω est un nombre réel bien défini, et qu'il vérifie

$$\lim_{\delta \rightarrow 0} \omega(\delta) = 0.$$

7. Montrer que pour tout $n \geq 0$, la famille $(B_{n,k})_{0 \leq k \leq n}$ est une base de l'espace des polynômes réels de degré inférieur ou égal à n .

8. Montrer que $B_n(f)$ est un polynôme de degré n , que l'on exprimera en fonction des $B_{n,k}$.

9. Montrer que

$$|B_n(f)(x) - f(x)| \leq \omega(\delta) + 2 \|f\|_\infty \mathbb{P}\left(\left|\frac{S_n}{n} - x\right| \geq \delta\right)$$

puis que

$$|B_n(f)(x) - f(x)| \leq \omega(\delta) + \frac{\|f\|_\infty}{2n\delta^2}.$$

Pour $a < b$ deux nombres réels quelconques, considérons maintenant l'espace $\mathcal{C}([a, b])$ des fonctions continues sur $[a, b]$ à valeurs réelles, que l'on munit de la norme $\|\cdot\|_\infty$ définie par $\|f\|_\infty = \sup_{[a, b]} |f|$ (on ne demande pas de vérifier qu'il s'agit bien d'une norme). On fixe $f \in \mathcal{C}([a, b])$.

10. Montrer que pour tout entier $n \geq 1$ il existe un polynôme $T_n(f)$ tel que

$$\|f - T_n(f)\|_\infty = d(f, \mathcal{P}_n),$$

où la distance est relative à la norme $\|\cdot\|_\infty$.

11. En utilisant les polynômes de Bernstein, montrer que $\|f - T_n(f)\|_\infty$ tend vers 0 quand n tend vers l'infini. Ainsi nous avons établi le théorème de Weierstrass.

12. On suppose dans cette question que f est lipschitzienne. Montrer qu'il existe une constante C telle que pour tout entier $n \geq 1$, $d(f, \mathcal{P}_n) \leq Cn^{-1/3}$.

13. Dans cette question, on se place sur $[a, b] = [-1, 1]$ et on considère la fonction valeur absolue $V : x \mapsto V(x)$. On admettra l'inégalité suivante dite des frères Markov (1890) : si P est un polynôme de degré au plus n , alors sur $[-1, 1]$ on a $\|P'\|_\infty \leq n^2 \|P\|_\infty$.

(a) Montrer qu'il existe un polynôme pair T_n^* tel que $d(V, \mathcal{P}_n) = \|V - T_n^*\|$.

(b) En minorant $|T_n^* - T_n^*(0) - V|$ sur un voisinage de 0, montrer qu'il existe une constante $c > 0$ telle que $d(V, \mathcal{P}_n) \geq cn^{-4}$.

Partie 4 : Théorème de Müntz.

Pour $\alpha \geq 0$ on se permettra de noter simplement x^α la fonction définie sur $[0, 1]$ par $x \mapsto x^\alpha$.

L'objet de cette partie est de démontrer le théorème suivant, dû à H. Müntz (1914) : soit $(\lambda_n)_{n \geq 0}$ une suite strictement croissante de nombres réels, telle que $\lambda_0 = 0$ et V le sous-espace vectoriel de $\mathcal{C}([0, 1])$ engendré par la famille de fonctions x^{λ_n} , $n \geq 0$. Alors V est dense dans $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$ si et seulement si la série $\sum_{n \geq 1} \frac{1}{\lambda_n}$ diverge.

Dans la suite on fixe $(\lambda_n)_{n \geq 0}$ une suite strictement croissante de nombres réels, telle que $\lambda_0 = 0$.

11. Pour $f \in \mathcal{C}([0, 1])$ on pose

$$\|f\|_2 = \left(\int_0^1 |f(t)|^2 dt \right)^{1/2}.$$

Vérifier que $\|\cdot\|_2$ définit une norme sur $\mathcal{C}([0, 1])$ associée à un produit scalaire que l'on explicitera et que l'on pourra noter $\langle \cdot | \cdot \rangle$ dans la suite.

12. Montrer que $\text{Vect}(\{x^m, m \in \mathbb{N}\})$ est dense dans $(\mathcal{C}([0, 1]), \|\cdot\|_2)$.
13. Pour $(\lambda_n)_{n \geq 0}$ comme ci-dessus, démontrer que la famille de fonctions (x^{λ_n}) est libre dans $\mathcal{C}([0, 1])$ (on pourra utiliser la question 4).
14. Soit $V_N = \text{Vect}(x^{\lambda_n}, 0 \leq n \leq N)$ et m un réel positif.

(a) Montrer que

$$d_2(x^m, V_N) = \frac{1}{\sqrt{2m+1}} \prod_{i=0}^N \frac{|\lambda_i - m|}{|\lambda_i + m + 1|},$$

où d_2 désigne la distance associée à la norme $\|\cdot\|_2$ sur $\mathcal{C}([0, 1])$.

(b) Montrer que si la suite (λ_n) est bornée alors $\lim_{N \rightarrow \infty} d_2(x^m, V_N) = 0$.

(c) Montrer que si (λ_n) est non-bornée et $m \notin \{\lambda_n, n \geq 0\}$ alors

$$\lim_{N \rightarrow \infty} d_2(x^m, V_N) = 0 \text{ si et seulement si } \sum_{n \geq 1} \frac{1}{\lambda_n} \text{ diverge.}$$

15. Montrer que V est dense dans $(\mathcal{C}([0, 1]), \|\cdot\|_2)$ si et seulement si la série $\sum_{n \geq 1} \frac{1}{\lambda_n}$ diverge.

16. Montrer que si V est dense dans $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$ alors la série $\sum_{n \geq 1} \frac{1}{\lambda_n}$ diverge.

17. (a) Prouver que si f et g sont de classe $\mathcal{C}^1$ sur $[0, 1]$ on a

$$\|f - g\|_\infty \leq |f(0) - g(0)| + \|f' - g'\|_2.$$

(b) Conclure que si la série $\sum_{n \geq 1} \frac{1}{\lambda_n}$ diverge alors V est dense dans $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$.

(c) Le théorème de Müntz reste-t'il vrai si l'on omet l'hypothèse $\lambda_0 = 0$?

◇

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2024

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet comporte deux problèmes qui peuvent être traités indépendamment. La notation tiendra largement compte de la clarté et de la précision des réponses.

Problème de la ruine du joueur

Deux individus A et B s'affrontent dans un jeu de hasard dont le principe est décrit ci-après. On note $k \in \mathbb{N} \setminus \{0, 1\}$ le montant global (en euro) mis en jeu. Au début de la partie, le joueur A dispose de $i \in \mathbb{N}$ euros. Le joueur B dispose donc de $k - i$ euros. Une tierce personne lance une pièce de monnaie : si elle tombe sur « face », le joueur A prend 1 euro au joueur B sinon, c'est le joueur B qui prend 1 euro au joueur A . Le jeu se termine lorsque l'un des deux joueurs est ruiné. Dans toute la suite du problème on notera $p \in]0, 1[$ la probabilité que la pièce tombe sur « face ».

L'objectif est de calculer (en fonction de p , k et i) la probabilité que le joueur A gagne la partie i.e., que le joueur B soit ruiné.

Pour tout $\ell \in \mathbb{N}$, on note $S_{i,k}^A(\ell)$ le montant dont dispose le joueur A au bout de ℓ lancers avec la condition initiale $S_{i,k}^A(0) = i$. Pour les 3 questions suivantes, on suppose que $k \geq 4$ et $i \in \{2, \dots, k-2\}$.

- 1) Quelles sont les valeurs prises par la variable aléatoire $S_{i,k}^A(1)$? Même question pour la variable aléatoire $S_{i,k}^A(2)$.
- 2) Déterminer la loi de la variable aléatoire $S_{i,k}^A(2)$.
- 3) Dans le cas où $p = 1/2$, quelle est l'espérance de $S_{i,k}^A(2)$?

On pose $\tau_{i,k}^A := \min \{ \ell \in \mathbb{N} \mid S_{i,k}^A(\ell) \in \{0, k\} \}$. La probabilité que le joueur A gagne est donc

$$w_{i,k}^A := \mathbb{P} \left(\left[S_{i,k}^A \left(\tau_{i,k}^A \right) = k \right] \right).$$

Pour tout $\ell \geq \tau_{i,k}^A$, on a $S_{i,k}^A(\ell) = S_{i,k}^A(\tau_{i,k}^A) \in \{0, k\}$. Autrement dit, dès lors que le joueur A a gagné ou perdu, le jeu s'arrête et la somme dont dispose le joueur A ne change plus.

- 4) Quelle est la valeur de la probabilité $w_{0,k}^A$ (i.e., la probabilité que le joueur A gagne s'il dispose de 0 euro en début de partie)?
- 5) Quelle est la valeur de la probabilité $w_{k,k}^A$?

Dans les questions 6) à 9), on se place dans le cas simple $i = 2$ et $k = 3$.

- 6) Calculer la probabilité pour que $\tau_{2,3}^A$ soit égal à 1.
- 7) Expliquer pourquoi $S_{2,3}^A(\tau_{2,3}^A) = 3$ si et seulement si $\tau_{2,3}^A$ est impair.
- 8) Montrer par récurrence que pour tout $m \in \mathbb{N}$,

$$\mathbb{P} \left(\left[\tau_{2,3}^A = 2m + 1 \right] \right) = p^{m+1} (1 - p)^m.$$

- 9) En déduire l'expression (en fonction de p) de $w_{2,3}^A$. Vous vérifierez que lorsque $p = 1/2$

(i.e., pour une pièce équilibrée), le joueur A a 2 chances sur 3 de remporter le jeu.

On se propose à présent de calculer la probabilité $w_{i,k}^A$ dans le cas général. Pour ce faire, on commence par faire l'étude de la suite $(a_{i,r})_{i \in \mathbb{N}}$ vérifiant pour $r > 0$ et tout $i \geq 1$ la relation de récurrence $a_{i+1,r} - a_{i,r} = r(a_{i,r} - a_{i-1,r})$.

10) Pour tout $i \geq 2$, donner l'expression de la fonction $g(i, r) > 0$ pour laquelle

$$a_{i,r} - a_{0,r} = g(i, r) (a_{1,r} - a_{0,r}).$$

(attention au cas $r = 1$).

11) Soit $n > 1$ un entier. En utilisant la question précédente, montrer que

$$a_{1,r} = \frac{a_{n,r} - a_{0,r}}{g(n, r)} + a_{0,r}.$$

Le calcul de $w_{i,k}^A$ est basé sur la remarque suivante. Pour tout $\ell \in \mathbb{N}$ et $(r_0, \dots, r_{\ell+1}) \in \{0, \dots, k\}^{\ell+2}$,

$$\begin{aligned} & \mathbb{P} \left(\left[S_{i,k}^A(\ell+1) = r_{\ell+1} \right] \mid \left[S_{i,k}^A(\ell) = r_\ell \right] \cap \dots \cap \left[S_{i,k}^A(0) = r_0 \right] \right) \\ &= \mathbb{P} \left(\left[S_{i,k}^A(\ell+1) = r_{\ell+1} \right] \mid \left[S_{i,k}^A(\ell) = r_\ell \right] \right). \end{aligned}$$

Ainsi, à partir du $(\ell+1)$ -ième lancer, la probabilité que le joueur A gagne la partie ne dépend que de la somme $S_{i,k}^A(\ell)$ dont il dispose après ℓ lancers de la pièce. Autrement dit pour tout $\ell \in \mathbb{N} \setminus \{0\}$ et $j \in \{0, \dots, k\}$,

$$\mathbb{P} \left(S_{i,k}^A(\tau_{i,k}^A) = k \mid S_{i,k}^A(\ell) = j \right) = \mathbb{P} \left(S_{j,k}^A(\tau_{j,k}^A) = k \right).$$

12) En utilisant l'égalité ci-dessus, justifier la relation de récurrence

$$w_{i,k}^A = p w_{i+1,k}^A + (1-p) w_{i-1,k}^A, \text{ pour tout } i \geq 1.$$

13) En utilisant les réponses aux questions 10), 11) et 12), donner l'expression (en fonction de p , k et i) de la probabilité $w_{i,k}^A$.

14) Vérifier que le résultat ci-dessus coïncide lorsque $i = 2$ et $k = 3$ à la réponse donnée à la question 9). Vous pourrez dans un premier temps montrer l'égalité ci-dessous qui est vraie pour tout $p \in]0, 1[\setminus \{1/2\}$

$$\frac{p^3 - (1-p)^3}{p^2 - (1-p)^2} = 1 - p(1-p).$$

Une autre méthode pour estimer la probabilité que le joueur A gagne est de faire l'étude du processus $\{S_{i,k}^A(\ell); \ell \in \mathbb{N}\}$. Comme nous l'avons remarqué précédemment, sa loi dépend uniquement des probabilités (appelées « probabilités de transition ») définies pour tout $\ell \in \mathbb{N}$ par

$$p_{r,s} := \mathbb{P}\left(\left[S_{i,k}^A(\ell+1) = s\right] \mid \left[S_{i,k}^A(\ell) = r\right]\right) = \mathbb{P}\left(\left[S_{i,k}^A(1) = s\right] \mid \left[S_{i,k}^A(0) = r\right]\right), \quad (r, s) \in \{0, \dots, k\}^2.$$

La matrice de transition $\mathbf{P}$ est la matrice de dimension $(k+1) \times (k+1)$ dont l'élément de la ligne u et de la colonne v avec $(u, v) \in \{1, \dots, k+1\}^2$ est $\mathbf{P}_{u,v} := p_{u-1, v-1}$.

- 15) Quelle est la première ligne de $\mathbf{P}$? La dernière ligne de $\mathbf{P}$?
- 16) Pour tout $u \in \{2, \dots, k\}$ donner les expressions, en fonction de p , de $\mathbf{P}_{u,u-1}$, $\mathbf{P}_{u,u+1}$ et $\mathbf{P}_{u,v}$ pour tout $v \notin \{u-1, u+1\}$.
- 17) On rappelle qu'une matrice stochastique est une matrice dont tous les éléments sont positifs ou nuls et dont la somme des éléments de chaque ligne est égal à 1. La matrice $\mathbf{P}$ est-elle une matrice stochastique?
- 18) Pour tout couple $(r, s) \in \{0, \dots, k\}^2$, donner l'expression, en fonction des probabilités $p_{r,t}$ et $p_{t,s}$ avec $t \in \{0, \dots, k\}$, de la probabilité

$$\mathbb{P}\left(\left[S_{i,k}^A(\ell+2) = s\right] \mid \left[S_{i,k}^A(\ell) = r\right]\right), \quad \ell \in \mathbb{N}.$$

- 19) Soit $(u, v) \in \{1, \dots, k+1\}^2$. A l'aide de la question précédente, expliciter la probabilité donné par l'élément $\mathbf{P}_{u,v}^2$ de la ligne u et de la colonne v de la matrice $\mathbf{P}^2$? De manière plus générale, pour $m \in \mathbb{N} \setminus \{0\}$, que représente l'élément $\mathbf{P}_{u,v}^m$?
- 20) La matrice $\mathbf{P}^m$ est-elle une matrice stochastique? (justifier votre réponse).

On admettra que lorsque $m \rightarrow \infty$, la matrice $\mathbf{P}^m$ converge vers la matrice stochastique $\mathbf{Q}$.

- 21) Comment interprétez vous les éléments de la $(k+1)$ -ième colonne de $\mathbf{Q}$? Justifier votre réponse.

Pour les questions suivantes, on considère le cas $k = 3$.

- 22) Donner la matrice de transition $\mathbf{P}$.
- 23) Pour tout $m \in \mathbb{N} \setminus \{0\}$, donner la première et la dernière ligne de la matrice $\mathbf{P}^m$.
- 24) On pose $q := 1 - p$. Dans le cas où $m \in \mathbb{N} \setminus \{0\}$ est pair, montrer par récurrence que la deuxième ligne de $\mathbf{P}^m$ est le vecteur ligne

$$\left(q \sum_{j=0}^{m/2-1} (pq)^j, (pq)^{m/2}, 0, p^2 \sum_{j=0}^{m/2-1} (pq)^j \right),$$

et que la troisième ligne est le vecteur ligne

$$\left(q^2 \sum_{j=0}^{m/2-1} (pq)^j, 0, (pq)^{m/2}, p \sum_{j=0}^{m/2-1} (pq)^j \right).$$

- 25) En utilisant la question précédente, donner la deuxième et la troisième ligne de la matrice $\mathbf{P}^m$ lorsque m est impair.
- 26) Montrer que lorsque $m \rightarrow \infty$, la matrice $\mathbf{P}^m$ converge vers la matrice $\mathbf{Q}$ que vous préciserez. Vous vérifierez également que $\mathbf{Q}$ est une matrice stochastique.
- 27) Avant de commencer le jeu, une tierce personne lance la pièce : si elle tombe sur « face », le joueur A reçoit la somme de 1 euro (et donc le joueur B reçoit 2 euros) sinon il reçoit 2 euros. Calculer la probabilité que le joueur A gagne en fonction de p .

Problème du char d'assaut allemand

Lors de la seconde guerre mondiale, les chars allemands étaient identifiables par un numéro de série visible sur le châssis. Les numéros de série étaient des entiers consécutifs et celui du premier char fabriqué était inconnu. En utilisant simplement les numéros de série des quelques chars capturés, les alliés souhaitaient connaître le nombre de chars d'assaut produits par les allemands.

L'énoncé mathématique de ce problème est le suivant. Soient $s \in \mathbb{N}$, $N \in \mathbb{N} \setminus \{0\}$ et $n \in \{1, \dots, N\}$. Les valeurs s et N sont inconnues. Dans l'ensemble $\{s+1, \dots, s+N\}$, on tire simultanément n éléments $x_1, \dots, x_n$. Les valeurs ordonnées correspondantes seront notées $x_{(1)} < \dots < x_{(n)}$. Le n -uplet $(x_{(1)}, \dots, x_{(n)})$ peut être vu comme la réalisation d'une variable aléatoire

$$(X_{(1)}, \dots, X_{(n)}) : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (E_n, \mathcal{P}(E_n)),$$

où $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace probablisé, $E_n = \{(i_1, \dots, i_n) \in \{1, \dots, N\}^n \mid i_1 < \dots < i_n\}$ et $\mathcal{P}(E_n)$ est la tribu des parties de E_n . En utilisant uniquement les observations $x_{(1)}, \dots, x_{(n)}$, il s'agit donc de proposer une estimation de l'entier N .

- 1) A quoi correspondent les entiers s , N et n dans le problème des chars d'assaut?
- 2) Quel est le cardinal de l'ensemble E_n ?
- 3) Quelle est la loi de la variable aléatoire $(X_{(1)}, \dots, X_{(n)})$? Autrement dit, quelle est la valeur de la probabilité $\mathbb{P}([(X_{(1)}, \dots, X_{(n)}) = (i_1, \dots, i_n)])$ où $(i_1, \dots, i_n) \in E_n$?

Soient $N \in \mathbb{N} \setminus \{0\}$, $n \in \{1, \dots, N\}$ et $m \in \{1, \dots, n\}$. Une variable aléatoire Y suit une loi **hypergéométrique inverse** de paramètres (N, n, m) (notation : $Y \sim \mathcal{H}(N, n, m)$) si pour tout $i \in \{m, \dots, N - n + m\}$

$$\mathbb{P}([Y = i]) := \frac{\binom{i-1}{m-1} \binom{N-i}{n-m}}{\binom{N}{n}}.$$

La somme des probabilités étant égale à 1, on a

$$\sum_{i=m}^{N-n+m} \binom{i-1}{m-1} \binom{N-i}{n-m} = \binom{N}{n}.$$

4) En utilisant l'égalité ci-dessus et la relation

$$i \binom{i-1}{m-1} = m \binom{i}{m},$$

montrer que si $Y \sim \mathcal{H}(N, n, m)$ alors $\mathbb{E}(Y) = m(N+1)/(n+1)$.

5) Montrer que si $Y \sim \mathcal{H}(N, n, m)$ alors

$$\text{Var}(Y) = \frac{m(N+1)}{(n+1)^2(n+2)} (N-n)(n-m+1).$$

Soit $k \in \{1, \dots, n\}$. On s'intéresse à la variable aléatoire $X_{(k)}$ qui est la k -ième composante de la variable aléatoire $(X_{(1)}, \dots, X_{(n)})$.

6) Quel est l'ensemble des valeurs prises par la variable aléatoire $X_{(k)}$?

7) Montrer que $X_{(k)} - s$ suit une loi hyper-géométrique inverse dont vous préciserez les paramètres.

8) Montrer que pour tout $1 \leq k < \ell \leq n$,

$$\hat{N}(n, k, \ell) := \frac{n+1}{\ell-k} (X_{(\ell)} - X_{(k)}) - 1,$$

est un estimateur sans biais de N (i.e., d'espérance N).

9) Quelle est la valeur de l'estimateur $\hat{N}(n, k, \ell)$ lorsque $n = N$ (i.e., lorsqu'on tire tous les éléments de $\{s+1, \dots, s+N\}$) ?

Afin de calculer la variance de l'estimateur ci-dessus, on s'intéresse à présent à la loi du couple aléatoire $(X_{(k)}, X_{(\ell)})$ avec $1 \leq k < \ell \leq n$.

Pour simplifier, on supposera que $s = 0$ dans toute la suite.

10) Justifier l'équivalence : l'événement $[X_{(k)} = i] \cap [X_{(\ell)} = j]$ est non vide si et seulement si $k \leq i$, $j - i \geq \ell - k$ et $N - j \geq n - \ell$. Vous pourrez vous aider du schéma ci-dessous.

$$\begin{array}{ccccccc} 1 & \cdots & i & \cdots & j & \cdots & N \\ \cdots x_{(1)} \cdots & x_{(k)} & \cdots & x_{(\ell)} & \cdots & x_{(n)} \cdots \end{array}$$

11) Que faut-il mettre à la place des points d'interrogation pour que

$$E^*(N, n, k, \ell) = E^* := \{(i, j) \in \mathbb{N}^2 \mid k \leq i \leq ? \text{ et } \ell - k + i \leq j \leq ?\},$$

soit l'ensemble des valeurs prises par le couple $(X_{(k)}, X_{(\ell)})$?

12) Pour tout $(i, j) \in E^*$, donner l'expression de la probabilité $\mathbb{P}([X_{(k)} = i] \cap [X_{(\ell)} = j])$.

13) En déduire que

$$\sum_{(i,j) \in E^*} \binom{i-1}{k-1} \binom{j-i-1}{\ell-k-1} \binom{N-j}{n-\ell} = \binom{N}{n}.$$

14) Montrer que pour tout $1 \leq k < \ell \leq n$,

$$\mathbb{E}(X_{(k)}(X_{(\ell)} - X_{(k)})) = \frac{\ell - k}{n - k + 1} \left((N + 1)\mathbb{E}(X_{(k)}) - \mathbb{E}(X_{(k)}^2) \right).$$

15) En déduire l'expression de la covariance entre $X_{(k)}$ et $X_{(\ell)}$ lorsque $1 \leq k < \ell \leq n$.

Vous devez donner une expression de la forme $\text{Cov}(X_{(k)}, X_{(\ell)}) = r(n, k, \ell)\text{Var}(X_{(k)})$, où $r(n, k, \ell)$ est à préciser.

16) Utiliser les résultats précédents pour montrer que

$$\text{Var}(\hat{N}(n, k, \ell)) = \frac{(N + 1)(N - n)}{(\ell - k)^2(n + 2)} [k(n - k + 1) + (\ell - 2k)(n - \ell + 1)].$$

Il reste maintenant à choisir les valeurs de k et ℓ avec $1 \leq k < \ell \leq n$ pour obtenir le « meilleur » estimateur de N . Il paraît naturel d'utiliser la plus grande observation pour estimer au mieux la valeur N . On prend donc $\ell = n$. Reste à choisir $k \in \{1, \dots, n - 1\}$ pour avoir un estimateur avec la plus petite variance $\sigma_n^2(k) := \text{Var}(\hat{N}(n, k, n))$ possible.

17) Après avoir fait l'étude de la fonction $k \in \{1, \dots, n - 1\} \mapsto \sigma_n^2(k)$, que proposez-vous comme choix pour k ?

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2025

Épreuve de mathématiques

Durée : 4h

L'épreuve est constituée de trois problèmes indépendants.

Rappel : Pour toutes les épreuves, les appareils électroniques (calculatrices, téléphones, traductrices), les casques, et les dictionnaires, ainsi que tout document, sont interdits.

Problème 1 : algèbre linéaire.

Dans ce problème on étudie l'existence de racines carrées réelles ou complexes pour les matrices $d \times d$ à coefficients réels, c'est-à-dire que l'on se donne $A \in M_d(\mathbb{R})$ et on recherche une matrice M dans $M_d(\mathbb{R})$ ou dans $M_d(\mathbb{C})$ telle que $M^2 = A$. On parlera alors respectivement de racine carrée réelle ou complexe de A .

Dans les questions 1 à 5 on s'intéresse au cas $d = 2$ et on considère ensuite le cas où la dimension d est quelconque.

1. Supposons que $A \in M_2(\mathbb{R})$ est diagonalisable sur $\mathbb{R}$ et à valeurs propres positives. Montrer que A admet une racine carrée dans $M_2(\mathbb{R})$.
2. Supposons que $A \in M_2(\mathbb{R})$ admet deux valeurs propres non nulles de signes opposés.
 - (a) Montrer que A admet une racine carrée dans $M_2(\mathbb{C})$.
 - (b) Est-ce que A admet une racine carrée dans $M_2(\mathbb{R})$?
3. Pour $\theta \in \mathbb{R}$, on pose $B_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. En s'appuyant sur une interprétation géométrique de la matrice B_θ , montrer qu'elle admet une racine carrée dans $M_2(\mathbb{R})$.
4. Supposons maintenant que $A \in M_2(\mathbb{R})$ est nilpotente et que $A \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
 - (a) Montrer que $A^2 = 0$.
 - (b) La matrice A admet-elle une racine carrée dans $M_2(\mathbb{C})$?
5. Supposons maintenant que $A \in M_2(\mathbb{R})$ admet une valeur propre réelle mais n'est pas diagonalisable sur $\mathbb{R}$.
 - (a) Montrer qu'il existe $\lambda \in \mathbb{R}$ et une matrice nilpotente non nulle N tels que $A = \lambda I + N$.
 - (b) En supposant $\lambda > 0$, montrer que A admet une racine carrée dans $M_2(\mathbb{R})$.
 - (c) On suppose maintenant $\lambda < 0$. La matrice A peut-elle être diagonalisable sur $\mathbb{C}$?
 - (d) Montrer que A n'admet pas de racine carrée dans $M_2(\mathbb{R})$.
6. On rappelle que $GL_d(\mathbb{K})$ désigne l'ensemble des matrices inversibles à coefficients dans $\mathbb{K}$. Soient A et B deux matrices de $M_d(\mathbb{R})$ telles qu'il existe $P \in GL_d(\mathbb{C})$ telle que $B = P^{-1}AP$. On pose $P = P_1 + iP_2$, avec $P_1, P_2 \in GL_d(\mathbb{R})$.
 - (a) Montrer qu'il existe $x \in \mathbb{R}$ tel que $\det(P_1 + xP_2) \neq 0$.
 - (b) Montrer que A et B sont semblables dans $M_d(\mathbb{R})$.
 - (c) Application : montrer que si une matrice $A \in M_2(\mathbb{R})$ admet une valeur propre complexe de la forme $\alpha + i\beta$ avec $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^*$, alors A est semblable sur $\mathbb{R}$ à la matrice $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$.
7. Soit $A \in M_d(\mathbb{R})$ une matrice diagonalisable sur $\mathbb{C}$. On suppose que toutes les valeurs propres négatives de A sont de multiplicité paire. Montrer que A admet une racine carrée dans $M_d(\mathbb{R})$.

Problème 2 : analyse.

Soit f une fonction continue sur $]0, +\infty[$. Sous réserve que cette expression ait un sens, on pose pour $x \in \mathbb{R}_+$

$$F(x) = \int_0^\infty \frac{tf(t)}{x^2 + t^2} dt.$$

L'objet de ce problème est d'étudier cette transformation et d'en déduire le calcul de certaines intégrales.

1. Dans cette question on considère la fonction f définie par $f(t) = \frac{\operatorname{Arctan} t}{t^2}$.

- (a) Montrer que $F(x)$ est bien définie pour tout $x \in]0, +\infty[$.
- (b) On pose $G(x) = \int_0^\infty \frac{\operatorname{Arctan}(xt)}{t(1+t^2)} dt$. Exprimer F en fonction de G .
- (c) Montrer avec précision que G est de classe C^1 sur $[0, +\infty[$.
- (d) Calculer G' et en déduire la valeur de $F(x)$ pour tout x .
- (e) Montrer que la fonction $t \mapsto \left(\frac{\operatorname{Arctan} t}{t}\right)^2$ est intégrable sur $[0, \infty[$ et utiliser ce qui précède pour déterminer la valeur de l'intégrale

$$\int_0^\infty \left(\frac{\operatorname{Arctan} t}{t}\right)^2 dt.$$

2. Dans cette question on considère la fonction f définie par $f(t) = \frac{\cos t}{t}$.

- (a) Montrer que $F(x)$ est bien définie pour tout $x \in]0, +\infty[$.
- (b) On pose H la fonction définie par $H(x) = xF(x)$. Montrer que la fonction H est bornée sur $\mathbb{R}_+$ et que $\lim_{x \rightarrow 0} H(x) = \frac{\pi}{2}$. (On pourra penser au changement de variable $t = ux$).
- (c) Démontrer que F est de classe C^2 sur $]0, \infty[$.
- (d) En admettant la relation (que l'on ne demande pas de démontrer)

$$\frac{\partial^2}{\partial t^2} \left(\frac{x}{x^2 + t^2} \right) + \frac{\partial^2}{\partial x^2} \left(\frac{x}{x^2 + t^2} \right) = 0,$$

établir que H est solution d'une équation différentielle linéaire du second ordre à coefficients constants.

- (e) En déduire l'expression de F .

Problème 3 : probabilités.

Dans tout ce problème, $(\Omega, \mathcal{A}, \mathbb{P})$ désigne un espace probabilisé sur lequel seront définies les différentes variables aléatoires considérées. L'espérance d'un évènement $A \in \mathcal{A}$ sera notée $\mathbb{E}(A)$ et sa probabilité sera notée $\mathbb{P}(A)$.

On pourra utiliser sans démonstration le lemme de Borel-Cantelli : si $(A_n)_{n \geq 1}$ est une suite d'évènements tels que la série $\sum_{n \geq 1} \mathbb{P}(A_n)$ converge, alors pour presque tout ω , on a $\omega \notin A_n$ pour n assez grand.

Soit Z une variable aléatoire réelle discrète. On note (M) la condition

$$(M) \quad \text{pour tout } \alpha \in \mathbb{R}_+, \mathbb{E}(e^{\alpha|Z|}) \text{ est finie.}$$

On admettra que si (M) est satisfaite alors l'espérance $\mathbb{E}(|Z|)$ est finie, et on définit alors

$$\psi_Z(\lambda) = \ln \left(\mathbb{E} \left(e^{\lambda(Z - \mathbb{E}(Z))} \right) \right).$$

1. Montrer que si la variable aléatoire Z est bornée, la condition (M) est satisfaite.
2. On suppose que Z suit une loi de Poisson de paramètre $\theta > 0$. Montrer que la condition (M) est satisfaite et déterminer la fonction ψ_Z .
3. On suppose que Z est une variable aléatoire satisfaisant la condition (M). Montrer que pour tout $\lambda \in \mathbb{R}_+$ et pour tout $t \in \mathbb{R}$ on a

$$\mathbb{P}(Z - \mathbb{E}(Z) \geq t) \leq e^{-(\lambda t - \psi_Z(\lambda))}.$$

4. Soit g la fonction définie sur $\mathbb{R}^+$ par

$$g(x) = \ln(\text{ch}(x)) - \frac{x^2}{2},$$

où ch désigne la fonction cosinus hyperbolique : $\text{ch}(x) = \frac{e^x + e^{-x}}{2}$.

- (a) Montrer que la fonction g est concave.
- (b) En déduire que pour tout $x \geq 0$ on a $\text{ch}(x) \leq e^{x^2/2}$.
5. Soient c un réel strictement positif et X une variable aléatoire réelle discrète satisfaisant $\mathbb{E}(X) = 0$ et $|X(\omega)| \leq c$ pour tout $\omega \in \Omega$.
 - (a) Montrer qu'il existe une variable aléatoire Y à valeurs dans $[0, 1]$ telle que $X = -cY + c(1 - Y)$, et déterminer $\mathbb{E}(Y)$.
 - (b) Démontrer que pour tout $\lambda \geq 0$, on a $e^{\lambda X} \leq Y e^{-c\lambda} + (1 - Y) e^{c\lambda}$. (On pourra utiliser un argument de convexité.)
 - (c) Conclure que pour tout $\lambda \geq 0$, on a $\mathbb{E}(e^{\lambda X}) \leq \text{ch}(c\lambda)$.
6. On considère maintenant une suite $(X_n)_{n \geq 1}$ de variables aléatoires mutuellement indépendantes telles que pour tout $n \geq 1$, on a $\mathbb{E}(X_n) = 0$ et $|X_n| \leq c_n$, où $c_n \in \mathbb{R}_+^*$. On pose S_n la variable aléatoire définie par $S_n = \sum_{k=1}^n X_k$.

(a) Montrer que pour tout $\lambda \geq 0$ on a $\mathbb{E}\left(e^{\lambda S_n}\right) \leq e^{\frac{\lambda^2 v_n}{2}}$, où l'on a noté $v_n = \sum_{k=1}^n c_k^2$.

(b) En déduire que pour tout $t \geq 0$ on a

$$\mathbb{P}(|S_n| \geq t) \leq 2e^{-\frac{t^2}{2v_n}}.$$

(c) On suppose que $c_n \leq Cn^\alpha$ pour une certaine constante $C > 0$ et un réel $\alpha \geq 1$. Montrer qu'il existe une constante $D \in \mathbb{R}$ telle que $v_n \leq Dn^{2\alpha+1}$.

(d) Conclure que si $\beta > \frac{\alpha+1}{2}$ est fixé, alors pour presque tout $\omega \in \Omega$, pour tout n assez grand on a $|S_n| \leq n^\beta$.

◇◇◇

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2025

Épreuve à option (A) : Mathématiques

Durée : 4h

Rappel : Pour toutes les épreuves, les appareils électroniques (calculatrices, téléphones, traductrices), les casques, et les dictionnaires, ainsi que tout document, sont interdits.

Ce problème porte sur la notion de matrice de Hadamard. Les quatre parties sont largement indépendantes, même si les résultats principaux des parties 1 et 2 seront utilisés dans les parties 3 et 4.

Partie 1 : Une démonstration du théorème spectral.

On munit $\mathbb{R}^n$ de sa structure euclidienne canonique, le produit scalaire est noté $\langle \cdot, \cdot \rangle$ et la norme euclidienne est notée $\|\cdot\|$. On identifiera si besoin $\mathbb{R}^n$ avec l'ensemble des vecteurs colonnes $M_{n,1}(\mathbb{R})$. Étant donnée une matrice $M \in M_n(\mathbb{R})$ on note $M_{i,j}$ le coefficient d'indice (i, j) de M et M^T la transposée de M .

Si besoin, on pourra utiliser sans démonstration le fait que l'application

$$M \mapsto |||M||| = \left(\sum_{i,j=1}^n M_{i,j}^2 \right)^{1/2}$$

définit une norme sur $M_n(\mathbb{R})$.

On rappelle que $O_n(\mathbb{R})$ désigne le groupe des matrices orthogonales de taille n , c'est à dire l'ensemble des matrices des isométries vectorielles.

L'objet de cette partie est de fournir une démonstration du théorème spectral, c'est à dire du fait que toute matrice symétrique est diagonalisable en base orthonormale. Soit $A = (A_{i,j})_{1 \leq i,j \leq n} \in M_n(\mathbb{R})$ une matrice symétrique.

1. Montrer l'équivalence entre les conditions :

- (i) Pour tout $X \in \mathbb{R}^n$ on a $\|AX\| = \|X\|$;
- (ii) Pour tous $(X, Y) \in (\mathbb{R}^n)^2$ on a $\langle AX, AY \rangle = \langle X, Y \rangle$;
- (iii) $A^T A = I_n$.

2. Montrer que $O_n(\mathbb{R})$ est un sous ensemble compact de $M_n(\mathbb{R})$.

3. On pose ϕ l'application de $O_n(\mathbb{R})$ à valeurs dans $\mathbb{R}^+$ définie par

$$\phi(P) = \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} ((P^{-1}AP)_{i,j})^2.$$

Justifier que $\phi(P) = 2 \sum_{1 \leq i < j \leq n} ((P^{-1}AP)_{i,j})^2$ et montrer que ϕ atteint son minimum sur $O_n(\mathbb{R})$.

4. Pour $\theta \in [0, 2\pi]$ et r, s des entiers appartenant à $\{1, \dots, n\}$ tels que $r < s$, on pose $Q = Q(r, s, \theta)$ la matrice telle que

$$Q_{r,r} = Q_{s,s} = \cos \theta, \quad Q_{r,s} = -Q_{s,r} = \sin \theta, \quad Q_{i,i} = 1 \text{ pour } i \notin \{r, s\} \text{ et } Q_{i,j} = 0 \text{ sinon.}$$

(a) Montrer que Q appartient à $O_n(\mathbb{R})$.

(b) On pose $B = Q^{-1}AQ$. Calculer $B_{i,j}$ en distinguant les cas suivants :

- $i, j \notin \{r, s\}$
- $(i \notin \{r, s\}, j = r)$ et $(i \notin \{r, s\}, j = s)$
- $(i, j) = (r, s)$.

- (c) En déduire que $\phi(Q) = \phi(I_n) + 2(B_{r,s}^2 - A_{r,s}^2)$.
5. Montrer que si la matrice A est telle que $\phi(I_n) > 0$, il existe $(r, s, \theta) \in \{1, \dots, n\}^2 \times [0, 2\pi]$ et Q comme précédemment tels que $\phi(Q) < \phi(I_n)$.
6. Utiliser ce qui précède pour démontrer que toute matrice symétrique est diagonalisable en base orthonormale.

Partie 2 : Inégalité de Hadamard

On rappelle qu'une matrice symétrique est dite positive si ses valeurs propres sont positives ou nulles, et définie-positive si celles-ci sont strictement positives. On rappelle également qu'une matrice symétrique est positive (resp. strictement positive) si pour tout $X \in \mathbb{R}^n$ différent de 0, on a $\langle AX, X \rangle \geq 0$ (resp. $\langle AX, X \rangle > 0$).

7. Démontrer l'inégalité arithmético-géométrique : pour tout $(x_1, \dots, x_n) \in (\mathbb{R}_+)^n$, on a

$$(x_1 \cdots x_n)^{1/n} \leq \frac{x_1 + \cdots + x_n}{n},$$

avec égalité si et seulement si $x_1 = \cdots = x_n$.

8. Montrer que si B est une matrice symétrique positive on a $\det(B) \leq \left(\frac{\text{tr}(B)}{n}\right)^n$ et caractériser le cas d'égalité.
9. L'objet de cette question est de démontrer que si S est une matrice symétrique définie positive, on a

$$\det(S) \leq \prod_{i=1}^n S_{i,i}.$$

- (a) Montrer que pour tout $1 \leq i \leq n$, on a $S_{i,i} > 0$.
- (b) On pose D la matrice diagonale telle que pour $1 \leq i \leq n$, $D_{i,i} = (S_{i,i})^{-1/2}$. Déterminer DSD et vérifier que DSD est symétrique et définie positive.
- (c) Démontrer l'inégalité indiquée, et caractériser le cas d'égalité.
10. Soit maintenant $A \in \text{GL}_n(\mathbb{R})$ une matrice inversible quelconque.
- (a) Montrer que la matrice $A^T A$ est symétrique définie positive.
- (b) En utilisant les questions précédentes, démontrer l'inégalité de Hadamard

$$|\det(A)| \leq \prod_{i=1}^n \|C_i\|,$$

où les C_i sont les vecteurs colonnes de A .

- (c) Montrer que cette inégalité est une égalité si et seulement si les colonnes de A sont orthogonales.

Partie 3 : Matrices de Hadamard

Soit n un entier supérieur ou égal à 2. On pose

$$\mathcal{E}_n = \{A \in M_n(\mathbb{R}) : \forall (i, j) \in \{1, \dots, n\}^2, |A_{i,j}| \leq 1\},$$

$$\mathcal{E}_n^+ = \{A \in M_n(\mathbb{R}) : \forall (i, j) \in \{1, \dots, n\}^2, |A_{i,j}| = 1\}$$

et $\mathcal{H}_n$ l'ensemble des matrices de $\mathcal{E}_n^+$ dont les colonnes sont orthogonales. Par définition, $\mathcal{H}_n$ est l'ensemble des matrices de Hadamard de taille n .

On s'intéresse au problème suivant : déterminer le maximum de la fonction déterminant sur $\mathcal{E}_n$. On note

$$m_n = \max\{\det(A), A \in \mathcal{E}_n\}.$$

11. Montrer que la fonction déterminant atteint son maximum sur $\mathcal{E}_n$, et que celui-ci est atteint sur $\mathcal{E}_n^+$.

(Indication : pour cela, étant donnée $A \in \mathcal{E}_n$, on pourra considérer pour (i_0, j_0) fixé la matrice $A + tE_{i_0, j_0}$, où $(E_{i,j})$ désigne la base canonique de $M_n(\mathbb{R})$)

12. Montrer que $m_n = \max\{|\det(A)|, A \in \mathcal{E}_n\}$.
13. Montrer que $m_n \leq n^{n/2}$ et qu'une matrice $A \in \mathcal{E}_n$ réalise la borne $|\det(A)| = n^{n/2}$ si et seulement si elle est de Hadamard.
14. Montrer que si n est impair alors $\mathcal{H}_n$ est vide. Montrer que $\mathcal{H}_2$ est non-vide.
15. Montrer que si $A \in \mathcal{E}_n^+$ alors $\det(A)$ est un entier pair.
16. En déduire que $m_3 = 4$.
17. Soit $n \geq 3$. Dans cette question on montre que si $\mathcal{H}_n$ est non-vide, alors n est un multiple de 4. Pour une matrice $A \in M_n(\mathbb{R})$, on note $C_i(A)$ le i^{e} vecteur colonne de A .
- (a) Montrer que si $\mathcal{H}_n \neq \emptyset$, il existe $H \in \mathcal{H}_n$ telle que $C_1(H) = (1, \dots, 1)$.
- (b) On prend H comme à la question précédente et on introduit les quantités suivantes :

- x le cardinal de l'ensemble des indices i tels que $H_{i,2} > 0$ et $H_{i,3} > 0$;
- y le cardinal de l'ensemble des indices i tels que $H_{i,2} > 0$ et $H_{i,3} < 0$;
- z le cardinal de l'ensemble des indices i tels que $H_{i,2} < 0$ et $H_{i,3} > 0$;
- t le cardinal de l'ensemble des indices i tels que $H_{i,2} < 0$ et $H_{i,3} < 0$

Écrire un système de quatre équations satisfaites par les variables x, y, z, t et en déduire que n est un multiple de 4.

18. Soit $H \in \mathcal{H}_n$ et soit $\tilde{H}$ la matrice définie par blocs par

$$\tilde{H} = \begin{pmatrix} H & H \\ H & -H \end{pmatrix}.$$

Montrer que $\tilde{H}$ appartient à $\mathcal{H}_{2n}$. En déduire que pour tout $k \geq 1$, l'ensemble des matrices de Hadamard de taille 2^k est non vide.

Jacques Hadamard a produit en 1893 des exemples de matrices de Hadamard de taille 12 et 20. La question de savoir si pour tout k il existe une matrice de Hadamard de taille $4k$ est ouverte depuis plus d'un siècle.

Partie 4 : Méthode probabiliste et minoration de m_n

Dans cette partie on se donne un entier $n \geq 2$ et on considère une matrice aléatoire notée M dont les coefficients $M_{i,j}$ sont des variables aléatoires indépendantes et identiquement distribuées, suivant une loi uniforme dans $\{-1, 1\}$. On admettra qu'il existe un espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$ sur lequel seront définies les diverses variables aléatoires considérées et l'espérance sera notée $\mathbb{E}(\cdot)$.

On rappelle la formule

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)},$$

où $\mathfrak{S}_n$ désigne le groupe symétrique (i.e. le groupe des bijections de $\{1, \dots, n\}$) et $\varepsilon(\sigma)$ est la signature de la permutation σ .

19. Pour toute fonction $f : \mathcal{E}_n^+ \rightarrow \mathbb{R}$, justifier que $\mathbb{E}(f(M))$ s'exprime sous la forme

$$\mathbb{E}(f(M)) = \frac{1}{N_n} \sum_{A \in \mathcal{E}_n^+} f(A),$$

où N_n est une quantité que l'on explicitera.

20. Montrer que $\mathbb{E}(\det(M)) = 0$.

21. (a) Soient σ et τ des permutations distinctes de $\mathfrak{S}_n$. Montrer que

$$\mathbb{E}(M_{1,\sigma(1)} \cdots M_{n,\sigma(n)} \cdot M_{1,\tau(1)} \cdots M_{n,\tau(n)}) = 0.$$

(b) En déduire que $\mathbb{E}(\det(M)^2) = \mathbb{E}(\det(MM^T)) = n!$

22. En déduire que $m_n \geq \sqrt{n!}$

23. Soit (u_n) la suite définie pour $n \geq 1$ par

$$u_n = \frac{n!}{n^n e^{-n} \sqrt{n}}.$$

L'objet de cette question est de montrer que la suite u_n converge vers une limite $\ell > 0$.

(a) Soit $v_n = \ln \frac{u_{n+1}}{u_n}$. Faire un développement limité à l'ordre 2 de v_n quand $n \rightarrow \infty$.

(b) Démontrer le résultat annoncé.

24. Montrer que pour n assez grand on a

$$m_n \geq \left(\frac{n}{e}\right)^{n/2}.$$

Montrer également que si θ est un réel tel que $\frac{1}{e} < \theta \leq 1$, alors

$$\mathbb{P}(|\det(M)| \geq (\theta n)^{n/2}) \xrightarrow{n \rightarrow \infty} 0.$$

On conjecture qu'il existe une constante $c > 0$ telle que pour tout n , $m_n \geq cn^{n/2}$. Terence Tao et Van Vu ont démontré en 2005 que pour tout $\delta > 0$, avec une grande probabilité, on a

$$e^{-\delta n} \leq \frac{|\det(M)|}{\sqrt{n!}} \leq e^{\delta n}.$$

◇

Banque d'Épreuves des Concours des Écoles d'Actuariat et Statistique

Session 2025

Épreuve à option (B) : Probabilités

Durée : 4h

Le sujet est composé d'un unique problème ayant pour objet l'estimation de quantiles extrêmes. Les questions peuvent être traitées indépendamment. Il est cependant conseillé de lire l'intégralité du sujet avant de commencer. La notation tiendra largement compte de la clarté et de la précision des réponses.

Notations et rappels

Pour l'ensemble de l'énoncé, on considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ où $\mathcal{A}$ est une tribu de parties de Ω et $\mathbb{P}$ une mesure de probabilité.

- ↪ Une **variable aléatoire réelle** est une fonction $X : \Omega \rightarrow \mathbb{R}$ telle que pour tout $x \in \mathbb{R}$, $[X \leq x] = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{A}$.
- ↪ Soit X une variable aléatoire. La fonction $F_X(\cdot)$ définie pour tout $x \in \mathbb{R}$ par $F_X(x) = \mathbb{P}([X \leq x])$ est la **fonction de répartition** de X . On admettra que cette fonction est nécessairement croissante et continue à droite.
- ↪ Un **vecteur aléatoire** de taille $n \geq 1$ est un n -uplet $(X_1, \dots, X_n)$ dont toutes les composantes sont des variables aléatoires réelles.
- ↪ Deux vecteurs aléatoires $(X_1, \dots, X_n)$ et $(Y_1, \dots, Y_n)$ sont dits **égaux en loi** si pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$, $\mathbb{P}([X_1 \leq x_1] \cap \dots \cap [X_n \leq x_n]) = \mathbb{P}([Y_1 \leq x_1] \cap \dots \cap [Y_n \leq x_n])$. L'égalité en loi est notée $(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (Y_1, \dots, Y_n)$.
- ↪ La loi de la variable aléatoire X est absolument continue sur $\mathbb{R}$ si la fonction de répartition $F_X(\cdot)$ est continue et dérivable (presque-partout) de dérivée $f_X(\cdot)$ qui s'intègre à 1. On a alors que pour toute fonction continue positive $g : \mathbb{R} \rightarrow [0, \infty[$,

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

- ↪ Soit $(X_n, n \in \mathbb{N})$ une suite de variables aléatoires. On dit que X_n **converge en probabilité** vers $c \in \mathbb{R}$ ($X_n \xrightarrow{\mathbb{P}} c$) si pour tout $\varepsilon > 0$, $\mathbb{P}(|X_n - c| \geq \varepsilon) \rightarrow 0$ lorsque $n \rightarrow \infty$. On admettra également que pour toute fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ continue, si $X_n \xrightarrow{\mathbb{P}} c$ alors $g(X_n) \xrightarrow{\mathbb{P}} g(c)$.
- ↪ Soit $(X_n, n \in \mathbb{N})$ une suite de variables aléatoires. On dit que X_n **converge en loi** vers une loi $\mathcal{D}$ ($X_n \xrightarrow{\mathcal{L}} \mathcal{D}$) si $\mathbb{P}(X_n \leq t) \rightarrow G(t)$ en tout point de continuité de la fonction de répartition $G(\cdot)$ de la loi $\mathcal{D}$.
- ↪ Soit $(X_n, n \in \mathbb{N})$ une suite de variables aléatoires que l'on supposera indépendantes et de même loi. Si ces variables aléatoires admettent une espérance égale à $\mu \in \mathbb{R}$, la **loi faible des grands nombres** assure que

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} \mu.$$

- ↪ Soit $(X_n, n \in \mathbb{N})$ une suite de variables aléatoires que l'on supposera indépendantes et de même loi. Si ces variables aléatoires admettent une espérance égale à $\mu \in \mathbb{R}$ et une

variance égale à $\sigma^2 > 0$, le **théorème de la limite centrale** assure que

$$\frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{i=1}^n X_i - \mu \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Présentation du problème

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire de fonction de répartition $F_X(\cdot)$. Le **quantile** d'ordre $u \in [0, 1]$ de la variable aléatoire X est donné par

$$Q_X(u) := \inf \{x \in \mathcal{X} \mid F_X(x) \geq u\},$$

où $\mathcal{X} \subset \mathbb{R}$ est l'ensemble des valeurs prises par la variable X . La fonction $Q_X(\cdot)$ est appelée **la fonction quantile** de X . Le **point terminal** de la loi de la variable aléatoire X est

$$\xi_X := \lim_{u \rightarrow 1-} Q_X(u) \in]-\infty, +\infty].$$

Il correspond à la plus grande valeur observable de X . Enfin, on admettra facilement que si $F_X(\cdot)$ est strictement croissante et continue (donc bijective) alors $Q_X(\cdot) = F_X^{-1}(u)$ où $F_X^{-1}(\cdot)$ est la bijection réciproque.

A l'aide des observations de $n \in \mathbb{N} \setminus \{0\}$ variables aléatoires $X_1, \dots, X_n$ indépendantes et de même fonction de répartition $F_X(\cdot)$, l'objectif du problème est celui de l'estimation du quantile $Q_X(1 - \alpha_n)$ où (α_n) est une suite dans $[0, 1]$ telle que $n(1 - \alpha_n) \rightarrow c \in [0, 1)$ lorsque $n \rightarrow \infty$. Un tel quantile est dit **extrême**.

1 Premier contact avec la fonction quantile

Soit Y une variable aléatoire de fonction de répartition

$$F_Y(x) = \begin{cases} 0 & \text{si } x \geq 0, \\ \exp(-1/x) & \text{si } x < 0 \end{cases}$$

- 1) Quelles sont les valeurs prises par la variable aléatoire Y ?
- 2) Donner l'expression de la fonction quantile $Q_Y(\cdot)$ de la variable aléatoire Y .
- 3) Quel est le point terminal de la loi de Y ?

On note U une variable aléatoire de loi uniforme sur $[0, 1]$.

- 4) Donner l'expression de sa fonction de répartition $F_U(\cdot)$ ainsi que de sa fonction quantile $Q_U(\cdot)$.

Soit Z une variable aléatoire de fonction de répartition $F_Z(\cdot)$ donnée par

$$F_Z(x) = \begin{cases} 0 & \text{si } x < 0, \\ x/2 & \text{si } x \in [0, 3/2[, \\ 1 & \text{si } x \geq 3/2. \end{cases}$$

- 5) En notant $Q_Z(\cdot)$ la fonction quantile de Z , donner la valeur de $Q_Z(1/2)$, de $Q_Z(3/4)$ et de $Q_Z(9/10)$.
- 6) Quel est le point terminal de la loi de Z ?

Dans les questions 7) à 9), on s'intéresse à quelques propriétés de la fonction quantile $Q_X(\cdot)$ d'une variable aléatoire quelconque X .

- 7) Montrer que si $Q_X(u) \leq x$ alors $F_X(x) \geq u$. La réciproque est-elle vraie?
- 8) En utilisant la question précédente, montrer que $Q_X(U) \stackrel{\mathcal{L}}{=} X$ où U est une variable aléatoire de loi uniforme sur $[0, 1]$.
- 9) En déduire que si X admet une espérance (i.e., si $\mathbb{E}(|X|) < \infty$) alors

$$\mathbb{E}(X) = \int_0^1 Q_X(u) du.$$

2 Une première tentative d'estimation

Une première idée pour estimer le quantile extrême $Q_X(1 - \alpha_n)$ est d'utiliser l'estimateur empirique de la fonction de répartition. Cet estimateur est donné par

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \text{ind}_{]-\infty, x]}(X_i) = \frac{\text{Card}(\{i \in \{1, \dots, n\} \mid X_i \leq x\})}{n}.$$

où $\text{ind}_A(\cdot)$ est la fonction indicatrice sur l'ensemble A . On propose alors d'estimer $Q_X(\cdot)$ par $\hat{Q}_{n,X}(\cdot) := \inf\{x \in \mathcal{X} \mid \hat{F}_n(x) \geq \cdot\}$. Nous verrons plus loin que cet estimateur fait appel à la notion de statistique d'ordre. L'objectif de la partie 2.1 est de se familiariser avec cette notion. L'étude de l'estimateur $\hat{Q}_{n,X}(\cdot)$ sera faite dans la partie 2.2.

2.1 Statistiques d'ordre

Les statistiques d'ordre $X_{1,n} \leq \dots \leq X_{n,n}$ sont obtenues en triant par ordre croissant les valeurs prises par $X_1, \dots, X_n$. Par exemple, $X_{1,n} : \Omega \rightarrow \mathbb{R}$ est la nouvelle variable aléatoire définie pour tout $\omega \in \Omega$ par $X_{1,n}(\omega) = \min(X_1(\omega), \dots, X_n(\omega))$.

- 10) Donner, en fonction de $F_X(\cdot)$, l'expression de la fonction de répartition de la variable aléatoire $X_{1,n}$. Même question pour la variable aléatoire $X_{n,n}$.
- 11) Si $F_X(\cdot)$ est la fonction de répartition d'une loi uniforme sur $[0, 1]$, donner en fonction de n l'expression de $\mathbb{E}(X_{n,n})$.
- 12) Si $F_X(t) = \exp(-1/t^2)$ pour $t > 0$ et $F_X(t) = 0$ pour $t \leq 0$, donner en fonction de n l'expression de $\mathbb{E}(X_{n,n})$. On rappelle que

$$\int_0^\infty u^{-1/2} \exp(-u) du = \sqrt{\pi}.$$

Dans les questions 13) à 16), on s'intéresse aux statistiques d'ordre $U_{1,n} \leq \dots \leq U_{n,n}$ d'un échantillon $U_1, \dots, U_n$ de n variables aléatoires indépendantes de loi uniforme sur $[0, 1]$.

- 13) Pour $k \in \{1, \dots, n\}$, montrer l'égalité des événements

$$[U_{k,n} \leq x] = \left[B_n(x) := \sum_{i=1}^n \text{ind}_{]x, \infty]}(U_i) \leq n - k \right].$$

- 14) Quelle est la loi de la variable aléatoire $B_n(x)$?

- 15) En déduire que

$$\mathbb{P}([U_{k,n} \leq x]) = \sum_{r=k}^n \binom{n}{r} [F(x)]^r [1 - F(x)]^{n-r}.$$

- 16) En déduire également que $\mathbb{P}([U_{k,n} \leq x]) = \mathbb{P}([U_{n-k+1,n} > 1 - x])$ pour tout $k \in \{1, \dots, n\}$.

2.2 Estimateur empirique de la fonction quantile

- 17) Pour tout $x \in \mathbb{R}$, calculer l'espérance $\mathbb{E}(\hat{F}_n(x))$ et la variance $\mathbb{V}(\hat{F}_n(x))$ de la fonction de répartition empirique.
- 18) En utilisant la loi faible des grands nombres, montrer que $\hat{F}_n(x)$ converge en probabilité vers une constante que vous préciserez.
- 19) En utilisant le théorème de la limite centrale, préciser vers quelle loi converge la suite de variables aléatoires $\sqrt{n}(\hat{F}_n(x) - F(x))$.

- 20) Quelle est la valeur de $\hat{F}_n(x)$ pour $x < X_{1,n}$? Même question pour $x \geq X_{n,n}$.
- 21) Pour $k \in \{1, \dots, n-1\}$, quelle est la valeur de $\hat{F}_n(x)$ lorsque $X_{k,n} \leq x < X_{k+1,n}$?
- 22) En déduire que pour tout $u \in]0, 1[$, $\hat{Q}_{n,X}(u) = X_{[nu],n}$ où $[x] = \min\{k \in \mathbb{Z} \mid x \leq k\}$.
- 23) Pour $n = 10$, les valeurs observées des variables aléatoires $X_1, \dots, X_n$ sont

$$\begin{aligned} x_1 = 7.6 \quad x_2 = 5.7 \quad x_3 = 9.7 \quad x_4 = 4.9 \quad x_5 = 2.8 \quad x_6 = 2 \quad x_7 = 7.5 \\ x_8 = 0.5 \quad x_9 = 7.3 \quad x_{10} = 4.3 \end{aligned}$$

Quelle est la valeur observée de la médiane empirique $\hat{Q}_{n,X}(1/2)$? Celle du premier quartile empirique $\hat{Q}_{n,X}(1/4)$?

- 24) Quelle est l'expression de l'estimateur $\hat{Q}_{n,X}(1 - \alpha_n)$ lorsque $\alpha_n < 1/n$? Qu'en déduisez vous quant à la pertinence de cet estimateur pour estimer un quantile extrême?

3 Estimation des quantiles d'une loi de Pareto

On souhaite à présent estimer les quantiles extrêmes d'une **loi de Pareto**. La fonction de répartition d'une variable aléatoire X de loi de Pareto de paramètres $\gamma > 0$ et $c > 0$ est

$$F_X(x) = \begin{cases} 0 & \text{si } x < c, \\ 1 - (x/c)^{-1/\gamma} & \text{si } x \geq c \end{cases}$$

Dans la partie 3.1, on va s'intéresser à une famille de loi plus générale : **les lois de type Pareto**. L'estimation du quantile extrême $Q_X(1 - \alpha_n)$ sera faite dans la partie 3.2.

3.1 Lois de type Pareto

On dit qu'une variable aléatoire X suit une **loi de type Pareto** si sa fonction de répartition est de la forme

$$F_X(x) = 1 - x^{-1/\gamma} L(x),$$

avec $\gamma > 0$ et $L(\cdot)$ une fonction à variations lentes c'est-à-dire une fonction positive telle que pour tout $t > 0$,

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1.$$

En particulier, les fonctions positives qui convergent vers une constante $c > 0$ en l'infini sont des fonctions à variations lentes. On admettra enfin que le point terminal ξ_X d'une loi de type Pareto est $+\infty$.

- 25) En utilisant la définition ci-dessus, montrer que la fonction $x \mapsto \ln(x)$ est une fonction à variations lentes.
- 26) Lorsque la fonction de répartition $F_X(\cdot)$ est strictement croissante et continue, montrer que pour tout $u \in]0, 1[$,
- $$Q_X(u) = (1 - u)^{-\gamma} [L(Q_X(u))]^\gamma.$$
- 27) Si $L(x) \rightarrow c > 0$ lorsque $x \rightarrow \infty$, donner, en fonction de $s > 0$ et $\gamma > 0$, la limite du rapport $Q_X(1 - s\alpha)/Q_X(1 - \alpha)$ lorsque $\alpha \rightarrow 0$.
- 28) Soit la variable aléatoire Y de fonction de répartition $F_Y(\cdot)$ telle que $F_Y(x) = 0$ si $x \leq 0$ et $F_Y(x) = \exp(-1/x^2)$ si $x > 0$. Quelle est la limite de $x^2(1 - F_Y(x))$ lorsque $x \rightarrow \infty$.
- 29) En déduire que cette variable aléatoire Y suit une loi de type Pareto (vous préciserez la valeur de γ ainsi que l'expression de la fonction à variations lentes).
- 30) Calculer l'espérance de Y .

3.2 Estimateurs de Hill et de Weissman

Dans cette partie, on suppose que X suit une loi de Pareto de paramètres $\gamma > 0$ et $c > 0$. La loi de Pareto est évidemment une loi de type Pareto correspondant au cas où la fonction à variations lentes $L(\cdot)$ est constante égale à c .

Dans toute la suite, on notera $U_{1,n} \leq \dots \leq U_{n,n}$ les statistiques d'ordre associées à des variables aléatoires indépendantes $U_1, \dots, U_n$ suivant chacune la loi uniforme sur $[0, 1]$. On admettra les résultats suivants.

- (R1) Soit un entier $n \geq 1$, on note $E_1, \dots, E_{n+1}$ des variables aléatoires indépendantes suivant une loi exponentielle standard c'est-à-dire telles que pour tout $i = 1, \dots, n+1$,

$$\mathbb{P}([E_i \leq x]) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \exp(-x) & \text{si } x \geq 0. \end{cases}$$

Les statistiques d'ordre $U_{1,n} \leq \dots \leq U_{n,n}$ vérifient l'égalité en loi

$$(U_{i,n}, i = 1, \dots, n) \stackrel{\mathcal{L}}{=} \left(\sum_{j=1}^i E_j / \sum_{j=1}^{n+1} E_j, i = 1, \dots, n \right).$$

- (R2) $(X_{1,n}, \dots, X_{n,n}) \stackrel{\mathcal{L}}{=} (Q_X(U_{1,n}), \dots, Q_X(U_{n,n})).$

- (R3) $(U_{i,n}, i = 1, \dots, n) \stackrel{\mathcal{L}}{=} (1 - U_{n-i+1,n}, i = 1, \dots, n).$

- (R4) Si $X_n \xrightarrow{\mathcal{L}} \mathcal{D}$ et si $\epsilon_n \rightarrow 0$ alors $\epsilon_n X_n \xrightarrow{\mathbb{P}} 0$.

La première étape pour estimer un quantile extrême d'une loi de Pareto est l'estimation du paramètre $\gamma > 0$. Pour ce faire, on utilise l'**estimateur de Hill** défini par

$$\hat{\gamma}_n(k_n) := \frac{1}{k_n} \sum_{i=1}^{k_n} \ln \left(\frac{X_{n-i+1,n}}{X_{n-k_n,n}} \right).$$

où (k_n) est une suite d'entiers dans l'ensemble $\{1, \dots, n-1\}$ telle que $k_n \rightarrow \infty$ et $n/k_n \rightarrow \infty$.

31) Montrer que $(\ln(X_{n-i+1,n}), i = 1, \dots, n) \stackrel{\mathcal{L}}{=} (c(1 - U_{n-i+1,n})^{-\gamma}, i = 1, \dots, n)$.

32) En utilisant le résultat (R3), en déduire que

$$\hat{\gamma}_n(k_n) \stackrel{\mathcal{L}}{=} \frac{\gamma}{k_n} \sum_{i=1}^{k_n} \ln \left[\left(\frac{U_{i,n}}{U_{k_n+1,n}} \right)^{-1} \right].$$

33) En utilisant le résultat (R1), montrer que

$$\left(\frac{U_{i,n}}{U_{k_n+1,n}}, i = 1, \dots, k_n \right) \stackrel{\mathcal{L}}{=} (U_{i,k_n}, i = 1, \dots, k_n).$$

34) En déduire que

$$\hat{\gamma}_n(k_n) \stackrel{\mathcal{L}}{=} \frac{\gamma}{k_n} \sum_{i=1}^{k_n} \ln(U_i^{-1}).$$

35) Si U est une variable aléatoire de loi uniforme sur $[0, 1]$, quelle est la loi de la variable aléatoire $\ln(1/U)$?

36) En déduire, en utilisant la loi faible des grands nombres et le théorème de la limite centrale, les convergences

$$\hat{\gamma}_n(k_n) \xrightarrow{\mathbb{P}} \gamma \text{ et } k_n^{1/2}(\hat{\gamma}_n(k_n) - \gamma) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^2),$$

où $\mathcal{N}(0, \gamma^2)$ est une loi normale d'espérance nulle et de variance égale à γ^2 .

On estime finalement le quantile extrême d'une loi de Pareto par

$$\hat{Q}_{n,X}^{(W)}(1 - \alpha_n) := X_{n-k_n,n} \left(\frac{n\alpha_n}{k_n} \right)^{-\hat{\gamma}_n(k_n)}.$$

Cet estimateur est connu sous le nom d'**estimateur de Weissman**.

37) Montrer que $X_{n-k_n,n} \stackrel{\mathcal{L}}{=} cU_{k_n+1,n}^{-\gamma}$.

38) En utilisant le résultat (R1) et la loi faible des grands nombres, montrer que

$$\frac{n}{k_n} U_{k_n+1,n} \xrightarrow{\mathbb{P}} 1.$$

39) En déduire que

$$X_{n-k_n,n} \left(\frac{n}{k_n} \right)^{-\gamma} \xrightarrow{\mathbb{P}} c.$$

40) En utilisant (R4), montrer que si $k_n^{-1/2} \ln(k_n/(n\alpha_n)) \rightarrow 0$ alors

$$\left(\frac{n\alpha_n}{k_n} \right)^{\gamma - \hat{\gamma}_n(k_n)} \xrightarrow{\mathbb{P}} 1.$$

41) Conclure sur la convergence en probabilité de $\hat{Q}_{n,X}^{(W)}(1 - \alpha_n)/Q_X(1 - \alpha_n)$.