Devoir Surveillé 5 - Avec morceau HEC
Séries, probabilités discrètes, applications linéaires en dimension finie

La calculatrice est interdite. Durée de l'épreuve : 4h.

1. Lire le sujet en entier et en détail avant de commencer.
2. Une importance particulière sera attribuée à la rigueur de la rédaction et des justifications, qui devront être concises mais précises.
3. Une présentation agréable sera appréciée : laissez une marge à gauche, la première page libre, et encadrez vos résultats. **L'absence de respect de cette consigne entraîne une minoration de la note.**

Exercice 1 (Questions de cours)

1. Rappeler le théorème du rang et sa preuve.
2. Convergence absolue des séries : résultat et preuve.
3. Rappeler les formules d'espérance et de variance pour les lois de Poissons et les lois géométriques, et les démontrer.

Exercice 2 (EML ECS)

Soit n un entier naturel supérieur ou égal à 2. On note $E = \mathbb{R}_n[x]$ et $\mathscr{B} = (1, x, \ldots, x^n)$ la base canonique de E. On note, pour tout polynôme P de E :

$$\varphi(P) = \frac{1}{n}x(1-x)P' + xP. \qquad \textit{(P' étant le polynôme dérivé de P)}$$

1. a) Montrer que φ est une application linéaire.
   b) Calculer $\varphi(x^n)$.
   c) Montrer que φ est un endomorphisme de E.
2. Déterminer la matrice A de φ dans la base $\mathscr{B}$. Préciser le rang de cette matrice.
3. a) L'endomorphisme φ est-il injectif ? Justifier votre réponse.
   b) Soit P un polynôme non nul de Ker (φ).
   Montrer que P admet 1 comme racine puis montrer que l'ordre de multiplicité de 1 est n.
   c) En déduire une base de Ker (φ).
4. On pose, pour tout k de $[\![0, n]\!]$: $P_k = x^k(1-x)^{n-k}$.
   a) Pour tout k de $[\![0, n]\!]$, calculer $\varphi(P_k)$.
   b) Montrer que la famille $(P_0, P_1, \ldots, P_n)$ est une base de E.
   c) Donner la matrice B de φ dans cette base.

Problème - morceau de sujet HEC

Dans ce problème, toutes les variables aléatoires sont supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. *Si* X *est une variable aléatoire, on note respectivement* $\mathrm{E}(X)$ *et* $\mathrm{Var}(X)$ *son espérance et sa variance, sous réserve d'existence.*

Le but de ce problème est de mettre en évidence quelques résultats asymptotiques liés au modèle du collectionneur de vignettes. Dans chaque paquet de céréales se trouve une vignette et il y a en tout des vignettes de n types différents, où n est un entier supérieur ou égal à 1 . Chacun des n types de vignettes se retrouve avec la même fréquence dans les paquets de céréales. Une collection est alors complète lorsqu'elle comporte n vignettes de types différents.

On modélise le nombre total de paquets de céréales qu'il est nécessaire d'acheter pour obtenir la collection complète de n vignettes de types différents par la variable aléatoire notée C_n.

On pose par convention $C_0 = 0$ et pour tout $i \in [\![1, n]\!]$, on note C_i le nombre d'achats de paquets de céréales nécessaires pour obtenir i vignettes de types différents.

De même, pour tout $i \in [\![1, n]\!]$, on pose $X_i = C_i - C_{i-1}$, qui représente le nombre d'achats supplémentaires de paquets de céréales qu'il est nécessaire d'effectuer pour obtenir une nouvelle vignette d'un type différent des $(i-1)$ vignettes de types différents déjà obtenues. Par convention, on pose $X_1 = C_1 = 1$.

On suppose que les variables aléatoires $X_1, \ldots, X_n$ sont mutuellement indépendantes. Enfin, on pose :

$$V_n = \frac{C_n}{n} - \ln(n).$$

Questions préliminaires

1. a) Montrer que :

$$C_n = X_1 + \cdots + X_n = \sum_{i=1}^{n} X_i.$$

b) Justifier que pour tout $i \in [\![1, n]\!]$, la variable aléatoire X_i suit la loi géométrique $\mathcal{G}\left(\frac{n-i+1}{n}\right)$.

Partie I

Pour tout entier $n \geq 1$, on pose :

$$H_n = \sum_{k=1}^{n} \frac{1}{k} \quad , \quad S_n = \sum_{k=1}^{n} \frac{1}{k^2}.$$

Nous allons démontrer la convergence de la suite $(S_n)_{n\geq 1}$ *et déterminer une valeur approchée de sa limite* S.

2. a) Montrer que pour tout entier $k \geq 2$, on a l'encadrement :

$$\frac{1}{k} - \frac{1}{k+1} \leq \frac{1}{k^2} \leq \frac{1}{k-1} - \frac{1}{k}.$$

b) Montrer que pour tout entier $n \geq 2$, on a :

$$\frac{3}{2} - \frac{1}{n+1} \leq S_n \leq 2 - \frac{1}{n}.$$

c) Montrer que la suite $(S_n)_{n\geq 1}$ est convergente et donner un encadrement de sa limite S.

d) Montrer que pour tout entier $n \geq 1$, on a l'encadrement :

$$\frac{1}{n+1} \leq S - S_n \leq \frac{1}{n}.$$

e) En déduire un programme `Python` qui permet d'obtenir une valeur approchée de S à 10^{-7}-près.

3. Pour tout entier $n \geq 1$, on pose $u_n = H_n - \ln(n)$.

   a) Montrer que pour tout entier $n \geq 1$, on a l'encadrement :

   $$\frac{1}{n+1} \leq \ln\left(\frac{n+1}{n}\right) \leq \frac{1}{n}.$$

   b) Montrer que pour tout entier $n \geq 1$, on a : $0 \leq u_n \leq 1$.

   c) Montrer que la suite $(u_n)_{n\geq 1}$ converge vers une certaine limite $\gamma \in [0,1]$.

4. Justifier l'existence de l'espérance de C_n et montrer que $E(C_n) = nH_n$.

5. Justifier l'existence de la variance de C_n et exprimer $\mathrm{Var}(C_n)$ en fonction de n, S_n et H_n.

6. a) Montrer que pour tout réel $a > 0$, on a :

   $$\mathbb{P}(|C_n - nH_n| \geq an) \leq \frac{S}{a^2}.$$

   b) Montrer que pour tout réel $c > 1$, on a :

   $$\mathbb{P}\left(\left|\frac{C_n}{n} - \ln(n)\right| \geq c\right) \leq \frac{S}{(c-1)^2}.$$

   c) Pour cette question, on suppose que $n = 10^6$.

   On donne l'approximation $\ln(n) \approx 13.816$.

   Montrer que :

   $$\mathbb{P}\left(\frac{C_n}{n} \in [7.81, 19.82]\right) \geq 0.92.$$

Partie III - la partie II n'est pas faisable en ECG1

Dans cette partie, on suppose que $n \geq 2$. On suppose de plus que les vignettes sont numérotées de 1 à n. Pour tout $i \in [\![1,n]\!]$ et tout entier $m \geq 1$, on note $A_{i,m}$ la variable aléatoire donnant le nombre de vignettes numérotées i obtenues dans les m premiers paquets de céréales achetés.

7. Justifier que pour tout $i \in [\![1,n]\!]$, la variable aléatoire $A_{i,m}$ suit la loi binomiale $\mathcal{B}\left(m, \frac{1}{n}\right)$.

8. a) Montrer que si $(E_i)_{i\geq 1}$ est une famille d'événements, alors pour tout entier $r \geq 1$:

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

   b) Montrer que :

   $$\mathbb{P}(C_n > m) \leq n\left(1 - \frac{1}{n}\right)^m \leq n\mathrm{e}^{-\frac{m}{n}}.$$

   c) **Il y a ici une erreur de conception de sujet. Admettez cette question.** Soit c un réel strictement positif. Montrer qu'on a la majoration :

   $$\mathbb{P}(C_n > cn\ln(n)) \leq n^{1-c}.$$

d) Montrer que pour tout réel $x > -\ln(n)$, on a :

$$\mathbb{P}(V_n > x) \leq \mathrm{e}^{-x}.$$

Dans la suite de cette partie, on introduit un modèle légèrement différent : le nombre N de paquets achetés est décrit par une variable aléatoire N de loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda > 0$. On cherche à calculer la probabilité de compléter, à partir des N vignettes obtenues, la collection de vignettes. On suppose toujours que les vignettes sont numérotées de 1 à n. On note $\widetilde{A}_i$ la variable aléatoire donnant le nombre de vignettes numérotées i obtenues dans les N paquets de céréales achetés.

9. a) Soit $p \in \mathbb{N}^*$. Justifier que pour tout $i \in [\![1, n]\!]$, la loi conditionnelle de $\widetilde{A}_i$ sachant l'événement $[N = p]$ est la loi binomiale $\mathcal{B}\left(p, \frac{1}{n}\right)$ et en déduire que la variable aléatoire $\tilde{A}_i$ suit une loi de Poisson dont on déterminera le paramètre.

b) Montrer que pour tout n-uplet $(k_1, \ldots, k_n)$ avec $k_i \in \mathbb{N}$ et $k_1 + \cdots + k_n = p$, on a:

$$\mathbb{P}_{[N=p]}\left(\left[\tilde{A}_1 = k_1\right] \cap \cdots \cap \left[\tilde{A}_n = k_n\right]\right) = \frac{p!}{k_1! \cdots k_n!} \times \frac{1}{n^p}.$$

c) Montrer que pour tout n-uplet $(k_1, \ldots, k_n)$ avec $k_i \in \mathbb{N}$, on a :

$$\mathbb{P}\left(\left[\tilde{A}_1 = k_1\right] \cap \cdots \cap \left[\tilde{A}_n = k_n\right]\right) = \prod_{i=1}^{n} \frac{\mathrm{e}^{-\frac{\lambda}{n}} \left(\frac{\lambda}{n}\right)^{k_i}}{k_i!}.$$

d) Montrer que les variables aléatoires $\tilde{A}_1, \ldots, \tilde{A}_n$ sont mutuellement indépendantes. Commenter en comparant avec le résultat de la question 15.(c).

10. Soit D_n l'événement « À l'issue des N achats de paquets de céréales, la collection de vignettes est complète ».

Montrer que :

$$\mathbb{P}(D_n) = \left(1 - \mathrm{e}^{-\frac{\lambda}{n}}\right)^n.$$

On admet le résultat suivant : pour tout $p \in \mathbb{N}^*$, on a $\mathbb{P}_{[N=p]}(D_n) = \mathbb{P}(C_n \leq p)$.

11. a) Montrer que :

$$\mathbb{P}(D_n) = \sum_{p=0}^{\infty} \mathbb{P}(C_n \leq p)\, \mathbb{P}(N = p).$$

b) On suppose maintenant que $\lambda > 1$.

Soit $a \in]\sqrt{\lambda}, \lambda[$. On pose $k_1 = \lfloor \lambda - a \rfloor$ et $k_2 = \lfloor \lambda + a \rfloor + 1$.

Montrer que :

$$\sum_{p=k_1}^{k_2} \mathbb{P}(N = p) \geq 1 - \frac{\lambda}{a^2}.$$

c) Montrer l'encadrement :

$$\mathbb{P}(C_n \leq k_1)\left(1 - \frac{\lambda}{a^2}\right) \leq \mathbb{P}(D_n) \leq \mathbb{P}(C_n \leq k_2) + \frac{\lambda}{a^2}.$$

12. Soit $(c_n)_{n \geq 1}$ une suite convergente de nombres réels.

Montrer que pour tout entier n suffisamment grand, on a l'encadrement :

$$\mathbb{P}\left(V_n \leq c_n - \frac{1}{n^{1/3}} - \frac{1}{n}\right)\left(1 - \frac{\ln(n) + c_n}{n^{1/3}}\right) \leq \left(1 - \frac{e^{-c_n}}{n}\right)^n \leq \mathbb{P}\left(V_n \leq c_n + \frac{1}{n^{1/3}} + \frac{1}{n}\right) + \frac{\ln(n) + c_n}{n^{1/3}}.$$

On pourra appliquer la question précédente avec $\lambda = n\ln(n) + c_n n$ et $a = n^{2/3}$.