AVERTISSEMENT : **Il est rappelé à tous les candidats que le programme officiel de l'épreuve est le programme de Mathématiques des classes préparatoires au concours d'admission du groupe Sciences sociales (B/L) de la section des lettres de l'Ecole normale supérieure , dites "Khagnes S".**

Toute résolution faisant appel à des résultats ne figurant pas explicitement à ce programme sera rejetée.

Objet du Problème :

Le sujet est un problème d'analyse et de probabilités dont le but est la détermination des lois de probabilité de certaines variables aléatoires discrètes ou à densité définies à partir d'une suite de variables aléatoires indépendantes de même loi de densité exponentielle.
Les candidats sont invités à lire attentivement le problème en entier et à l'aborder dans son intégralité, quitte à utiliser des résultats fournis par l'énoncé même s'ils n'ont pas su les démontrer (en indiquant clairement quels résultats ils ont admis).
Les 3 parties sont indépendantes, sauf la question **II.3.b)** *qui utilise les résultats de la partie* **I**.

Notations :

- *Pour deux entiers naturels p et $q \geqslant p$, on note $[\![p,q]\!] = [p,q] \cap \mathbb{N}$, c'est-à-dire l'ensemble des entiers naturels compris, au sens large, entre p et q.*
- *Dans tout le problème, on considère un univers Ω muni d'une probabilité $\mathbf{P}$ et des variables aléatoires définies sur Ω, discrètes ou à valeurs dans $\mathbb{R}_+$, à densité continue sur $\mathbb{R}_+$. Dans ce cas, on ne s'intéressera qu'à la restriction à $\mathbb{R}_+$ de leur densité ou de leur fonction de répartition.*
- *Pour deux événements A et B, on désigne par $\mathbf{P}(A \mid B)$ la probabilité de A sachant B.*

Partie I

1° Soit f une fonction de classe C^∞ sur un intervalle ouvert I de $\mathbb{R}$ contenant 0. Prouver par récurrence sur p que

$$(\forall\, x \in I)\ (\forall\, p \in \mathbb{N}) \qquad f(x) = \sum_{k=0}^{p} \left(f^{(k)}(0)\, \frac{x^k}{k!} \right) + \int_0^x f^{(p+1)}(t)\, \frac{(x-t)^p}{p!}\, dt \tag{1}$$

où $f^{(k)}$ désigne la dérivée k-ième de f.

Pour $x \in [0,1[$, on pose $h(x) = \ln(1-x)$.

2° Calculer la dérivée p-ième de h.

3° Soit $x \in [0,1[$ fixé. Déterminer

$$\lim_{p \to +\infty} \left(\int_0^x h^{(p+1)}(t)\, \frac{(x-t)^p}{p!}\, \mathrm{d}t \right) .$$

On pourra commencer par étudier les variations de la fonction $\varphi(t) = \dfrac{t-x}{t-1}$ sur l'intervalle $[0,x]$.

4° En déduire que

$$(\forall\, x \in [0,1[) \qquad \ln(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k} . \tag{2}$$

5° Montrer que

$$(\forall x \in [0,1[) \qquad x + (1-x)\ln(1-x) = \sum_{k=1}^{\infty} \frac{x^{k+1}}{k(k+1)} \,. \tag{3}$$

Partie II

1° Soit X une variable aléatoire réelle à densité continue sur $\mathbb{R}_+$, qu'on peut interpréter comme la durée de vie d'un phénomène aléatoire, ayant la propriété suivante : la durée de vie résiduelle à chaque instant suit la même loi de probabilité que X.

a) Montrer que cette condition s'écrit :

$$(\forall s \in \mathbb{R}_+)\,(\forall t \in \mathbb{R}_+) \quad G(t+s) = G(t)\,G(s) \tag{4}$$

où l'on a posé, pour tout $t \in \mathbb{R}_+$, $G(t) = P(X > t)$.

b) Vérifier que la propriété (4) ci-dessus est satisfaite lorsque X suit une loi exponentielle sur $\mathbb{R}_+$.

Dans toute la suite du problème, on considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes et de même loi de probabilité exponentielle sur $\mathbb{R}_+$ de paramètre $\alpha > 0$, c'est-à-dire de densité

$$g(t) = \alpha e^{-\alpha t} \qquad \textit{pour} \quad t \in \mathbb{R}_+ \,.$$

On note F la fonction de répartition associée, c'est-à-dire la fonction définie sur $\mathbb{R}_+$ par

$$F(x) = P(X_0 \leqslant x) \,.$$

Dans la suite de cette partie, on considère la variable aléatoire N à valeurs dans $\mathbb{N}^*$ égale au plus petit indice $n \in \mathbb{N}^*$ tel que $X_n > X_0$.

2° **a)** Soit $x_0 \in \mathbb{R}_+^*$ et $n \in \mathbb{N}^*$. Exprimer $P(N = n \mid X_0 = x_0)$ en fonction de $p_0 = e^{-\alpha x_0}$.

b) En déduire que la loi de probabilité de N est

$$(\forall n \in \mathbb{N}^*) \qquad P(N = n) = \frac{1}{n\,(n+1)} \,.$$

c) La variable aléatoire N a-t-elle une espérance ?

3° **a)** Soit $x \in \mathbb{R}_+^*$ et $n \in \mathbb{N}^*$. Calculer $P(X_N \leqslant x \mid N = n)$.

b) Déterminer la fonction de répartition Ψ de X_N et sa densité ψ.
On pourra utiliser la partie I.

Partie III

Dans toute cette partie, m désigne un entier naturel non nul fixé. On range les variables aléatoires $X_0, X_1, \dots X_m$ par ordre *décroissant*, pour obtenir une nouvelle suite de variables aléatoires notées $Y_0, Y_1, \dots Y_m$. Ainsi, on a

$$Y_0 = \text{Max}\{X_k\,,\ 0 \leqslant k \leqslant m\} \quad \text{et} \quad Y_m = \text{Min}\{X_k\,,\ 0 \leqslant k \leqslant m\}\,.$$

1° Déterminer la fonction de répartition Φ_0 de Y_0 ainsi que sa densité φ_0.

2° Déterminer la fonction de répartition Φ_m de Y_m ainsi que sa densité φ_m.

3° On désigne par $\widehat{N}$ la variable aléatoire à valeurs dans $\mathbb{N}^*$ égale au plus petit indice $n \in \mathbb{N}^*$ tel que $X_{m+n} > Y_0$.

a) Soit $n \in \mathbb{N}$. Montrer que

$$P(\widehat{N} > n) = \frac{m+1}{m+n+1}\,.$$

b) En déduire la loi de la variable aléatoire $\widehat{N}$.

c) La variable aléatoire $\widehat{N}$ a-t-elle une espérance ?

4° **a)** Prouver que, pour tout $k \in [\![1, m]\!]$ et $(y_0, y_1, \dots, y_k) \in (\mathbb{R}_+)^{k+1}$, on a :

$$P(X_0 - X_1 > y_0,\ X_1 - X_2 > y_1,\ \dots,\ X_{k-1} - X_k > y_{k-1},\ X_k > y_k) = \frac{1}{(k+1)!}\exp\left[-\alpha\sum_{j=0}^{k}(j+1)\,y_j\right]. \tag{5}$$

On pourra procéder par récurrence sur k en exprimant la probabilité cherchée en fonction de

$$P(X_0 - X_1 > y_0,\ X_1 - X_2 > y_1,\ \dots,\ X_{k-2} - X_{k-1} > y_{k-2},\ X_{k-1} > y_{k-1} + t)\,,$$

pour t dans un domaine à préciser.

On admet que, pour tout $k \in \mathbb{N}^*$ et $(y_0, y_1, \dots, y_k) \in (\mathbb{R}_+)^{k+1}$, on a par symétrie :

$$P(Y_0 - Y_1 > y_0,\ Y_1 - Y_2 > y_1,\ \dots,\ Y_{k-1} - Y_k > y_{k-1},\ Y_k > y_k) = (k+1)!\ P(X_0 - X_1 > y_0,\ X_1 - X_2 > y_1,\ \dots,\ X_{k-1} - X_k > y_{k-1},\ X_k > y_k)\,. \tag{6}$$

b) En choisissant des valeurs particulières de $y_0, \dots, y_k$, prouver que, pour $k \in [\![1, m]\!]$, la loi de probabilité de la variable aléatoire $Y_{k-1} - Y_k$ a pour densité sur $\mathbb{R}_+$ la fonction

$$t \longrightarrow \alpha\, k\, e^{-\alpha k t}$$

et montrer que les variables aléatoires $Y_0 - Y_1$, $Y_1 - Y_2$, $\dots$, $Y_{m-1} - Y_m$ et Y_m sont indépendantes.

c) Montrer que, pour $k \in [\![1, m]\!]$, la variable aléatoire $Y_0 - Y_k$ a pour fonction de répartition sur $\mathbb{R}_+$

$$F_k(y) = [1 - e^{-\alpha y}]^k \tag{7}$$

et déterminer sa densité f_k.
On pourra encore procéder par récurrence sur k.

5° On désigne par $\widetilde{N}$ la variable aléatoire à valeurs dans $\mathbb{N}^*$ égale au plus petit indice $n \in \mathbb{N}^*$ tel que $X_{m+n} \notin [Y_m, Y_0]$, c'est-à-dire tel que

$$X_{m+n} < \mathrm{Min}\{X_k \,/\, 0 \leqslant k \leqslant m\} \quad \text{ou} \quad X_{m+n} > \mathrm{Max}\{X_k \,/\, 0 \leqslant k \leqslant m\}\,.$$

a) Soit $n \in \mathbb{N}^*$, $x \in \mathbb{R}_+^*$ et $z \in \mathbb{R}_+^*$. Calculer

$$P(\, x \leqslant X_{m+k} \leqslant x + z\,,\ \forall\, k \in [\![1,n]\!]\,)\,.$$

b) Soit $n \in \mathbb{N}^*$ et $z \in \mathbb{R}_+^*$. Montrer que

$$P(\, Y_m \leqslant X_{m+k} \leqslant Y_m + z\,,\ \forall\, k \in [\![1,n]\!]\,) = \frac{m+1}{m+n+1}\,[1 - e^{-\alpha z}]^n\,. \tag{8}$$

c) Soit $n \in \mathbb{N}$. Montrer que,

$$P(\widetilde{N} > n) = \frac{m(m+1)}{m+n+1} \int_0^{+\infty} \alpha\,[1 - e^{-\alpha z}]^{m+n-1}\, e^{-\alpha z}\,\mathrm{d}z \tag{9}$$

et en déduire $P(\widetilde{N} > n)$.

d) En déduire la loi de probabilité de $\widetilde{N}$.
La variable aléatoire $\widetilde{N}$ a-t-elle une espérance ?

6° Soit $k \in [\![0,m]\!]$.

a) Pour $z \in \mathbb{R}_+^*$, déterminer $P(Y_k \leqslant z)$.
On pourra exprimer l'événement $(Y_k \leqslant z)$ à l'aide des événements $(X_j \leqslant z)$, pour obtenir une expression de $P(Y_k \leqslant z)$ faisant intervenir un symbole $\sum$, qu'on ne cherchera pas à calculer explicitement.

b) En déduire que la densité de la loi de probabilité de Y_k est la fonction définie sur $\mathbb{R}_+$ par

$$\varphi_k(z) = \alpha\,(k+1)\,C_{m+1}^{k+1}\,[1 - e^{-\alpha z}]^{m-k}\, e^{-\alpha(k+1)z}\,. \tag{10}$$

c) Les variables aléatoires $(Y_k)_{0 \leqslant k \leqslant m}$ sont-elles indépendantes ?