DS9

Problème

On considère la fonction φ définie sur $]-\infty, 1]$ par :

$$\forall x \in]-\infty, 1], \quad \varphi(x) = \begin{cases} x + (1-x)\ln(1-x) & \text{si } x < 1 \\ 1 & \text{si } x = 1 \end{cases}$$

Partie A : Étude de la fonction φ

1. Montrer que la fonction φ est continue sur $]-\infty, 1]$.

2. ***a)*** Justifier que φ est de classe $\mathcal{C}^1$ sur $]-\infty, 1[$ et calculer, pour tout $x \in]-\infty, 1[$, $\varphi'(x)$.

b) En déduire les variations de φ sur $]-\infty, 1]$.

c) La fonction φ est-elle dérivable en 1 ?

3. Calculer la limite de φ en $-\infty$.

4. Tracer l'allure de la courbe représentative de φ en soignant le tracé aux voisinages de 0 et 1.

5. ***a)*** Calculer, à l'aide d'une intégration par parties, $\lim\limits_{A \to 0} \displaystyle\int_A^1 t\ln(t)\,dt$. On admet que $\displaystyle\int_0^1 t\ln(t)\,dt$ est une intégrale impropre en 0 **convergente** et que $\displaystyle\int_0^1 t\ln(t)\,dt = \lim_{A\to 0}\int_A^1 t\ln(t)\,dt$.

b) En déduire : $\displaystyle\int_0^1 \varphi(x)\,dx = \frac{1}{4}$.

Partie B : Étude de deux séries

Soit x un réel appartenant à $[0, 1[$.

6. ***a)*** Vérifier, pour tout n de $\mathbb{N}^*$ et tout t de $[0, x]$: $\displaystyle\frac{1}{1-t} - \sum_{k=0}^{n-1} t^k = \frac{t^n}{1-t}$.

b) En déduire, pour tout n de $\mathbb{N}^*$: $\displaystyle -\ln(1-x) - \sum_{k=1}^{n} \frac{x^k}{k} = \int_0^x \frac{t^n}{1-t}\,dt$.

7. Montrer, pour tout n de $\mathbb{N}^*$: $\displaystyle 0 \leqslant \int_0^x \frac{t^n}{1-t}\,dt \leqslant \frac{1}{(n+1)(1-x)}$.

En déduire la limite de $\displaystyle\int_0^x \frac{t^n}{1-t}\,dt$ lorsque l'entier n tend vers $+\infty$.

8. Montrer alors que la série $\displaystyle\sum_{n\geqslant 1} \frac{x^n}{n}$ converge et que l'on a : $\displaystyle\sum_{n=1}^{+\infty} \frac{x^n}{n} = -\ln(1-x)$.

9. ***a)*** Déterminer deux réels a et b tels que : $\forall n \in \mathbb{N}^*$, $\displaystyle\frac{1}{n(n+1)} = \frac{a}{n} + \frac{b}{n+1}$.

b) En déduire que la série $\sum\limits_{n\geqslant 1} \dfrac{x^{n+1}}{n\,(n+1)}$ converge et que l'on a : $\sum\limits_{n=1}^{+\infty} \dfrac{x^{n+1}}{n\,(n+1)} \;=\; \varphi(x)$.

10. Montrer que la série $\sum\limits_{n\geqslant 1} \dfrac{1}{n\,(n+1)}$ converge et que l'on a encore : $\sum\limits_{n=1}^{+\infty} \dfrac{1}{n\,(n+1)} \;=\; \varphi(1)$.

Partie C : Application en probabilité

Dans cette partie, toutes les variables aléatoires sont supposées définies sur un même espace probabilisé noté $(\Omega, \mathscr{A}, \mathbb{P})$.

On considère une urne contenant initialement une boule bleue et une boule rouge. On procède à des tirages successifs d'une boule au hasard selon le protocole suivant :

- si on obtient une boule bleue, on la remet dans l'urne et on ajoute une boule bleue supplémentaire ;
- si on obtient une boule rouge, on la remet dans l'urne et on arrête l'expérience.

On suppose que toutes les boules sont indiscernables au toucher et on admet que l'expérience s'arrête avec une probabilité égale à 1. On note N la variable aléatoire égale au nombre de boules présentes dans l'urne à la fin de l'expérience.

11. ***a)*** Montrer soigneusement : $\forall n \in \mathbb{N} \setminus \{0,1\}, \mathbb{P}\big([N = n]\big) = \dfrac{1}{n\,(n-1)}$.

b) La variable aléatoire N admet-elle une espérance ?

12. Recopier et compléter les lignes incomplètes de la fonction **Python** suivante de façon à ce qu'elle renvoie une simulation de la variable aléatoire N.

```
def simuleN():
    b = 1 # b désigne le nombre de boules bleues dans l'urne
    while rd.random() < .........
        b = b+1
    return .........
```

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes et de même loi. On suppose que, pour tout n de $\mathbb{N}^*$, les variables aléatoires $X_1, \ldots, X_n$ et N sont mutuellement indépendantes. On note F la fonction de répartition commune aux variables aléatoires X_n pour n appartenant à $\mathbb{N}^*$. On définit la variable aléatoire $T = \max(X_1, \ldots, X_N)$, ce qui signifie :

$$\forall \omega \in \Omega,\ T(\omega) \;=\; \max\big(X_1(\omega), \ldots, X_{N(\omega)}(\omega)\big)$$

Ainsi par exemple, si N prend la valeur 3, alors $T = \max(X_1, X_2, X_3)$; si N prend la valeur 5, alors $T = \max(X_1, X_2, X_3, X_4, X_5)$; etc.

13. ***a)*** Montrer : $\forall x \in \mathbb{R}, \forall n \in \mathbb{N} \setminus \{0,1\}, \mathbb{P}_{[N=n]}\big([T \leqslant x]\big) \;=\; \big(F(x)\big)^n$.

b) En déduire : $\forall x \in \mathbb{R}, \mathbb{P}\big([T \leqslant x]\big) \;=\; \varphi\big(F(x)\big)$.

14. On suppose **dans cette question uniquement** que, pour tout n de $\mathbb{N}^*$, la variable aléatoire X_n suit la loi uniforme sur $[0,1]$.

a) On rappelle que l'instruction `rd.random(d)` renvoie un tableau de taille `d` où les coefficients sont des réalisations de variables aléatoires indépendantes suivant la loi uniforme sur $[0,1]$.

Écrire une fonction **Python** nommée `simuleT()` qui renvoie une simulation de la variable aléatoire T.

b) On considère la fonction **Python** suivante :

```
def mystere():
    m = np.zeros(3)
    for k in range(3):
        s = np.zeros(1000)
        for j in range(1000):
            s[j] = simuleT()
        m[k] = np.mean(s)
    return m
```

À son appel, on obtient :

```
ans  =
     0.7474646     0.7577248     0.7470916
```

Que renvoie la fonction `mystere` ? Que peut-on conjecturer sur la variable aléatoire T ?

c) Montrer : $\forall x \in \mathbb{R},\ \mathbb{P}\big([T \leqslant x]\big) = \begin{cases} 0 & \text{si } x < 0 \\ \varphi(x) & \text{si } x \in [0,1[\\ 1 & \text{si } x \geqslant 1 \end{cases}$

d) En déduire que T est une variable aléatoire à densité et que la fonction

$$g : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ \varphi'(x) & \text{si } 0 \leqslant x < 1 \\ 0 & \text{si } x \geqslant 1 \end{cases}$$

est une densité de T.

e) Montrer, à l'aide d'une intégration par parties, que T admet une espérance et calculer $\mathbb{E}(T)$.
On admettra pour cela que $\displaystyle\int_0^1 t\,\varphi'(t)\,dt = \lim_{B\to 1} \int_0^B t\,\varphi'(t)\,dt$

15. On suppose **dans cette question uniquement** que, pour tout n de $\mathbb{N}^*$, la variable aléatoire X_n suit la loi exponentielle de paramètre λ avec $\lambda \in \mathbb{R}_+^*$.

a) Rappeler une expression de la fonction de répartition d'une variable aléatoire suivant la loi exponentielle de paramètre λ.

b) En déduire une expression de la fonction de répartition de T.

c) Montrer que T est une variable aléatoire à densité et qu'une densité de T est la fonction :

$$g : x \mapsto \begin{cases} 0 & \text{si } x < 0 \\ \lambda^2\, x\, \mathrm{e}^{-\lambda\, x} & \text{si } x \geqslant 0 \end{cases}$$

d) Justifier que T admet une espérance et que l'on a : $\mathbb{E}(T) = \lambda\ \mathbb{E}\big(X_1^2\big)$.
En déduire la valeur de $\mathbb{E}(T)$.

Exercice

Soient $m \in \mathbb{R}$ et $\sigma > 0$. Soit $\alpha \in]0,1[$.

1. On considère une variable aléatoire $X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right)$. On note $\varphi_{m,\sigma}$ la densité usuelle de X. Rappeler l'expression de $\varphi_{m,\sigma}(x)$ pour tout $x \in \mathbb{R}$.

2. On considère à partir de maintenant que le paramètre σ est connu mais que le paramètre m est inconnu. On cherche à estimer le paramètre m par la méthode du maximum de vraisemblance.
Soit $n \geqslant 2$ et soit $(X_1, \ldots, X_n)$ un n-échantillon de X.
Soit $(x_1, \ldots, x_n)$ une observation de ce n-échantillon. On définit la fonction de vraisemblance

$$L : a \mapsto \prod_{k=1}^{n} \varphi_{a,\sigma}(x_k)$$

a) Soit $a \in \mathbb{R}$. Calculer $L(a)$ puis $\ln(L(a))$.

b) On note $h : a \mapsto \ln(L(a))$. Montrer que la fonction h admet un unique maximum w sur $\mathbb{R}$ et exprimer w en fonction de $x_1, \ldots, x_n$ et n.

c) Que dire du réel w pour la fonction L ?

d) En déduire l'estimateur du maximum de vraisemblance pour le paramètre inconnu m.

3. *a)* On note $\overline{X_n} = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$. Montrer que $\overline{X_n} \hookrightarrow \mathcal{N}\left(m, \dfrac{\sigma^2}{n}\right)$.

b) En déduire que, pour tout $\varepsilon > 0$: $\mathbb{P}\left(\left[\overline{X_n} - \varepsilon \leqslant m \leqslant \overline{X_n} + \varepsilon\right]\right) \geqslant 1 - \dfrac{\sigma^2}{n\varepsilon^2}$.

c) Déduire de la question précédente un intervalle de confiance de m au niveau de confiance $1 - \alpha$.

d) On note $N(\alpha)$ le plus petit entier naturel n vérifiant : $\dfrac{\sigma}{\sqrt{\alpha n}} \leqslant 10^{-2}$. Calculer $N(\alpha)$ puis calculer $\lim\limits_{\alpha \to 0} N(\alpha)$.